Advances of matrix–analytic methods in risk modelling

Peralta Gutierrez, Oscar

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Advances of matrix–analytic methods in risk modelling

Oscar Peralta–Gutiérrez

Kongens Lyngby 2018

Technical University of Denmark
Department of Applied Mathematics and Computer Science
Richard Petersens Plads, building 324,
2800 Kongens Lyngby, Denmark
Phone +45 4525 3031
compute@compute.dtu.dk
www.compute.dtu.dk

Summary (English)

This work is concerned with the study of matrix–analytic methods with novel applications to the area of risk theory.

First, we review some topics of Applied Probability such as phase–type distributions, matrix–exponential distributions, Markovian arrival processes, Rational arrival processes, fluid flow processes and risk models. With these tools in hand, we propose a method to approximate the probability of ruin of any Cramér–Lundberg process using the theory of phase–type distributions, providing an error bound for such an approximation. With the goal of studying risk models with dependencies, we construct a class of bivariate distributions with given phase–type–distributed marginals and given Pearson's correlation coefficient, which are later used to define different kinds of dependent Sparre–Andersen processes. Later on, we give an explicit formula for the probability of Parisian and cumulative Parisian ruin for a class of risk processes which are based on the theory of fluid flow processes. Next, we study some excursion properties of spectrally negative Lévy processes whenever they are inspected at an independent matrix–exponential time. Finally, inspired by the generalisation of the Markovian arrival process to the Rational arrival process, we construct a novel generalisation of the fluid flow process without Brownian components and study its first passage probabilities.

Summary (Danish)

Dette arbejde omhnadler studiet af matrix analytiske metoder med nye applikationer inden for risikoteori.

Først gennemgår vi nogle emner af anvendt sandsynlighed såsom fasetype fordelinger, matrix–eksponentielle fordelinger, Markovske ankomstprocesser, rationelle ankomstprocesser, fluid flow processer og risikomodeller. Med disse værktøjer til rådighed, foreslår vi en metode til at tilnærme sandsynligheden for ruin i enhver Cramér–Lundberg proces ved hjælp teori om fasetype fordelinger, og giver en begrænsning af fejlen i en sådan approksimation. Med henblik på at studere risikomodeller med afhængigheder konstruerer vi en klasse af bivariate fordelinger med givne fasetypefordelte marginaler og givet Pearson korrelationskoefficient, som senere bruges til at definere forskellige typer af afhængige Sparre–Andersen processer. Derefter giver vi en eksplicit formel for sandsynligheden for parisisk og kumulativ parisisk ruin for en klasse af risikoprocesser, som er baseret på teori om fluid flow processer. Dernæst studerer vi nogle ekskursionsegenskaber af spektralt negative Lévy processer, når de inspiceres ved et uafhængig matrix–eksponentiel tidspunkt. Inspireret af generaliseringen af Markovske ankomstprocessen til den rationelle ankomstproces, konstruerer vi en ny generalisering af fluid flow processen uden Brownske komponenter og studerer dens første passage sandsynligheder.

Preface

This thesis was developed at the Section of Statistics and Data Analysis of the Department of Applied Mathematics and Computer Science at the Technical University of Denmark (DTU), in partial fulfilment of the requirements for acquiring the Ph.D. degree in Applied Mathematics.

This thesis, which is presented as a monograph, deals with the study of matrix–analytic methods and its use in risk theory.

This Ph.D. project was supervised by Professor Bo Friis Nielsen and co–supervised by Professor Mogens Bladt, researcher at KU (Department of Mathematical Sciences).

Lyngby, 30-September-2018

Oscar Peralta–Gutiérrez

Acknowledgements

I met Professor Mogens Bladt 7 years ago. I was being taught several aspects of risk theory during a BSc course at UNAM, but it was the elegance of one single proof which got stuck in my head for a while. It was about how to compute the probability of ruin of a Cramér–Lundberg process with phase–type distributed claims. Since then, I became fascinated with the field of matrix–analytic methods, or at that point, what I understood about it. Professor Mogens Bladt was kind enough to guide me through this field and others, and introducing me to his colleague, Professor Bo Friis Nielsen, at a later stage. When I finished my MSc studies, both of them encouraged me to pursue a Ph.D. at the Technical University of Denmark (DTU) under their supervision, doing research in the field that I liked the most. Of course I accepted. After three years, I would like to thank them for all the effort they put into my development as a researcher. Throughout my Ph.D. studies, they supported me continuously and helped me shaping my skills, for which I will be forever grateful. Their advice and expertise have always been invaluable to me. I learnt several things under their mentoring, both on an academic and on a personal level.

I would also like to express my deep gratitude to to Dr Leonardo Rojas–Nandayapa, who hosted me at the University of Liverpool for 1 week, and to Profesor Nigel Bean and Dr Giang Nguyen, who hosted me at The University of Adelaide for 6 months. These stays led me to realize that this field is much more vast than I initially imagined, which makes this academic journey even more exciting.

I thank DTU for providing me a healthy work environment. Thanks to all my friends and coworkers in Denmark, in particular I would like to thank Jes-

per Fink Andersen for helping me translate to danish some passages of this manuscript.

I am grateful for the financial support from the National Council of Science and Technology in México (CONACYT) during my Ph.D. studies.

Finally, I dedicate this thesis to all the people that I love and who love me. To those who are now in a better place. To those whose absence I suffered the most. To those who have tried their best to support me. This would have been impossible if it not were because of you. I promise everything will be worth it.

Abbreviations and notation

Matrices are denoted by bold capital letters and vectors by bold small letters. Special matrices and vectors are:

$\boldsymbol{I}$	an identity matrix appropiate dimensions
$\mathbf{0}$	a matrix of 0's of appropiate dimensions
$\boldsymbol{e}$	a column vector of 1's of appropiate dimension
$\boldsymbol{e}_i$	an unitary column vector with its i–th component equal to 1

For a pair of events A and B, a pair of random variables X and Y, a distribution G, a σ–algebra $\mathcal{F}$, and a collection of random variables $\{X_s\}_s$, we denote by:

$\mathbb{P}(A)$	the probability of the event A
$\mathbb{P}(A \mid B)$	the conditional probability of A given the event B
$\mathbb{E}(X)$	the expectation of X
$\mathrm{Var}(X)$	the variance of X
$X \perp Y$	the random variables X and Y are independent
$X \sim G$	the random variable X follows the distribution G
$\mathrm{b}\mathcal{F}$	the set of $\mathcal{F}$–measurable bounded functions
$\mathbb{P}(A \mid \mathcal{F})$	the conditional probability given $\mathcal{F}$
$\mathbb{E}(X \mid \mathcal{F})$	the conditional expectation of X given $\mathcal{F}$
$\sigma(\{X_s\}_s)$	the σ–algebra generated by $\{X_s\}_s$
$\mathbb{P}(A \mid \{X_s\}_s)$	the conditional probability given $\sigma(\{X_s\}_s)$
$\mathbb{E}(X \mid \{X_s\}_s)$	the conditional expectation given $\sigma(\{X_s\}_s)$

For matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, we denote by:

$e^{\boldsymbol{A}}$ or $\exp(\boldsymbol{A})$	the matrix exponential of $\boldsymbol{A}$
$\boldsymbol{A}'$	the transpose matrix of $\boldsymbol{A}$
$\boldsymbol{A} \otimes \boldsymbol{B}$	the Kronecker product of $\boldsymbol{A}$ and $\boldsymbol{B}$
$\boldsymbol{A} \oplus \boldsymbol{B}$	the Kronecker sum of $\boldsymbol{A}$ and $\boldsymbol{B}$
$\|\boldsymbol{A}\|_{\max}$	the max–norm of $\boldsymbol{A}$
$\mathrm{sp}(\boldsymbol{A})$	the spectra or collection of eigenvalues of $\boldsymbol{A}$
$\mathrm{dev}(\boldsymbol{A})$	the dominant eigenvalue of $\boldsymbol{A}$
$\lvert\boldsymbol{A}\rvert$	the entrywise–absolute–value matrix of $\boldsymbol{A}$

For a row or column vector $\boldsymbol{a}$, we denote by:

$\mathrm{diag}(\boldsymbol{a})$	the diagonal matrix constituted by the elements of $\boldsymbol{a}$
$(\boldsymbol{a})_i$	the i–th element of $\boldsymbol{a}$

For $z \in \mathbb{C}$, we denote by:

δ_z	the Dirac measure at z
$\mathrm{Re}(z)$	the real part of z
$\mathrm{Im}(z)$	the imaginary part of z

For a vector subspace or topological space X and a collection of $\{x_s\}_s$ with $x_s \in X$, we denote by:

$\dim(X)$	the dimension of X
$\mathrm{span}(\{x_s\}_s)$	the vector subspace generated by $\{x_s\}_s$
$\mathbb{B}(X)$	the Borel σ–algebra of X

For a functions f and g with domain $\mathbb{R}$ and $t \in \mathbb{R}$, we denote by:

$f(t^-)$	the limit $\lim_{s \uparrow t} f(s)$
$f(t^+)$	the limit $\lim_{s \downarrow t} f(s)$
f'	the derivative of f
$f * g$	the convolution of f and g
f^{*n}	the n–th convolution of f
$f \star g$	the Mellin convolution of f and g

Miscellaneous notation:

δ_{ij}	the Kronecker delta function
$o(h)$	a function f such that $f(h)/h \to 0$ as $h \downarrow 0$
$O(n^{-k})$	a function f such that for some $M > 0$, $f(n)/n^{-k} \leq M$ as $h \downarrow 0$
$M^{n \times m}(\mathbb{R})$	the set of $n \times m$–dimensional matrices with real entries

Finally, we use the following abbreviations throughout the text:

$\mathrm{Exp}(\lambda)$	the exponential distribution with mean $1/\lambda$
$\mathrm{Geo}(1-p)$	the geometric distribution with mass function $p^x(1-p)$ $(x \in \mathbb{Z}_+)$
PDMP	a piecewise deterministic Markov process
PH	a phase–type distribution
MPH^*	a Kulkarni's multivariate phase–type distribution
ME	a matrix–exponential distribution
MAP	a Markovian arrival process
RAP	a Rational arrival process
FRAP	a Fluid RAP

Contents

Chapter 1

Introduction

The behaviour of the capital of an insurance company over time can be modelled in many ways. Perhaps the most elementary continuous–time model is the process $\{R_t\}_{t\geq 0}$ with

$$R_t = u + pt + \sum_{i=1}^{N_t} U_i, \quad t \geq 0,$$

where $u \geq 0$, $p > 0$, $\{N_t\}_{t\geq 0}$ is a point process and $\{U_i\}_{i\geq 1}$ is a collection of nonnegative random variables. Indeed, at time 0 the insurance company has some fixed initial capital u, and over time, it continuously collects the premium paid by its customers at a rate p. Accidents to its customers occur according to the point process $\{N_t\}_{t\geq 0}$, and for each one of those incidents, the insurance company is required to cover each claim of size U_i at each arrival point. This mathematical framework was initially laid out by Filip Lundberg in the 1900's, whose results were later republished and studied in more detail by the Harald Cramér in the 1930's. These efforts ended up kickstarting the entire field of risk theory. One of the most interesting problems in risk theory is computing the probability that the insurance company will go bankrupt, that is, the probability that $\{R_t\}_{t\geq 0}$ eventually downcrosses 0. This is commonly referred as the ruin problem and is a non-trivial matter for most risk processes, since said probability depends on the assumptions made about the random behaviour of the stream of claim arrivals and the claim sizes themselves.

The exponential distribution is the most basic distribution with support in $[0, \infty)$, widely used in both simple and complex stochastic models. Its memoryless property is perhaps one of its main advantages, which is equivalent to the property of having a constant hazard rate. However, as much as we would like to convince ourselves that the exponential distribution arises naturally in real life situations, there are instances in which this is simply not true. One of the first mathematicians who tackled this matter was Anger K. Erlang, who found the exponential distribution to be inadequate when studying the telephone traffic for the Copenhagen Telephone Company in the 1900's. His idea was to construct a class of random variables linked to the lifetime of a hidden system with a fixed number of sequential stages. In order to keep the analysis of such a class of random variables simple, he proposed the completion time of each stage to be exponentially distributed, and all of the stages to be identically distributed. Nowadays, this class of random variables is said to follow an Erlang distribution. Half a century later, Arne Jensen proposed to generalise Erlang's idea of hidden sequential stages into a hidden Markovian system instead. More precisely, he considered a finite state Markov jump process which eventually gets absorbed into some state and studied the random time it took the process to get absorbed. The distribution of such a random variable is called phase–type (PH). Thus, a PH—distributed random variable has an underlying process which has exponentially distributed holding times of varying parameters. The study of PH distributions heavily relies on matrix computations and on the probabilistic interpretation of its underlying process. Popularity of PH distributions rose in the 1970's thanks to Marcel F. Neuts and co—authors, who laid out the theoretical basis of what became to be known as the matrix—analytic method. Over time, PH distributions became a highly regarded tool for modelling in Applied Probability because of their computational tractability and flexibility. The algebraic generalisation of the PH distribution is known as the matrix—exponential (ME) distribution. Even though most of the functions associated to an ME distribution (i.e. density function, distribution function, Laplace transform) are usually stated in terms of a matrix product just as in the case of PH distributions, ME distributed random variables do not necessarily have an underlying Markov jump process. Thus, ME distributions need to be analysed by either purely algebraic means or through more general underlying processes. Closely related to PH distributions are the Markovian arrival process (MAP) and the fluid flow process, both of which fall within the matrix—analytic method framework. The MAP is a point process which has an underlying Markov jump process, while the fluid flow process is an continuous time process which collects rewards from the holding times of an underlying Markov jump process. Just as in the case of PH and ME distributions, MAPs and fluid flow processes have been extensively used in stochastic modelling, especially in queueing theory and related areas. Analogous to the relation between PH and ME distributions, an algebraic extension of the MAP exists and is called the Rational arrival process (RAP).

The use of matrix—analytic methods in risk theory is not new. In fact, a considerable portion of the literature in risk theory is devoted to the use of models whose components (i.e. stream of claim arrivals or claim sizes) are of the matrix—analytic kind. In this thesis we explore further advances of the matrix—analytic method in risk theory. In particular, our approach to the ruin problem is threefold: by revisiting classic risk models from a novel perspective, by applying classic techniques to modern risk models, and by constructing new models from scratch. The structure and contents of this thesis are explained in detail next.

In Chapter 2 we lay a common ground of the machinery needed for the analysis in further chapters. More specifically, we review some Markovian models (Markov jump processes and piecewise deterministic Markov processes), probability distributions of the matrix–analytic kind (phase–type and matrix–exponential distributions), point processes of the matrix–analytic kind (Markovian arrival processes and Rational arrival processes), fluid flow processes with or without Brownian components, and some examples of risk models used in the literature. This chapter is intended to provide the reader a rigorous self–contained review of the material previously described. No novel results are presented in this chapter, though some of the proofs are different from the ones found in the literature. Discussions on the historical development of the topics are included as the theory unfolds.

Chapters 3 - 7 contain novel results. Because of this, their structure is somewhat different to the structure of Chapter 2. To keep a steady flow of the information presented, we postpone most discussions until the end of each chapter, where we provide a brief summary of our findings, their advantages and disadvantages, and their comparison with existing results in the literature. The novel results developed in this thesis are the following.

In Chapter 3 we revisit the ruin problem for the Cramér–Lundberg process. In particular, we provide a method for approximating the probability of ruin of a Cramér–Lundberg process with arbitrary claim size distribution. Such a method is based on the use of phase–type distributions, in particular, the Erlang distribution. Moreover, we provide bounds for the error of such an approximation.

In Chapter 4 we construct a bivariate distribution with given phase–type distributed marginals and given Pearson's correlation coefficient. Such a construction is based on the use of order statistics, and unlike most bivariate phase–type distributions, has an explicit density function. Additionally, we use such a bivariate distribution to construct Sparre–Andersen processes with different kinds of dependencies within their components.

In Chapter 5 we study the problem of Parisian ruin for a considerably large class of risk processes which are based in the theory of fluid flow processes. Parisian ruin is inspired in the idea that at each downcrossing of level 0, the regulatory agency for insurance companies gives them the chance to become "unruined" within some predetermined period of time (called clock) before declaring them in default. A related concept is cumulative Parisian ruin, in which the insurance company is declared to be in default if and only if the total time spent below 0 is greater than some predetermined period of time (also called clock). Using matrix–analytic methods, we provide a closed form formula for the probability of Parisian and cumulative parisian ruin in the case the clocks are phase–type distributed. Moreover, we discuss how an erlangization technique can be applied in order to approximate the probability of Parisian and cumulative Parisian ruin with deterministic clocks.

In Chapter 6 we revisit the Wiener–Hopf factorisation of a spectrally negative Lévy process. Such a factorisation characterises the behaviour of a spectrally negative Lévy process inspected at an independent exponentially distributed time. Using this classic result and functional calculus, we are able to characterise the behaviour of a spectrally negative Lévy process inspected at a matrix–exponential time. Moreover, we compute the probability of cumulative Parisian ruin for spectrally negative Lévy processes in the case the clock follows a matrix–exponential distribution; this result is based on the use of functional calculus too.

In Chapter 7 we construct a new stochastic model as follows. Just as ME distributions generalise PH distributions, and RAPs generalise MAPs, we construct a generalisation of the fluid flow process without Brownian components. We coin this process the Fluid RAP (FRAP). We precisely define the FRAP via its underlying process, which is very much inspired in the underlying process of a RAP. Later on, using novel techniques we show that the first passage probabilities of the FRAP are indeed an algebraic generalisation of the ones from the fluid flow processes.

Finally, in Chapter 8 we briefly present some research perspectives that emerge from the present work.

An appendix containing basic information about the Kronecker product and the Kronecker sum, holomorphic functional calculus for matrices, and the Richardson approximation formula is provided at the end of the thesis.

Some of the results presented in this thesis have been the subject of the following research articles:

- **O Peralta**, L Rojas-Nandayapa, W Xie, H Yao
  Approximation of ruin probabilities via erlangized scale mixtures. Insurance: Mathematics and Economics 78 (2018): 136-156.

  This paper contains part of the results in Chapter 3.

- M Bladt, BF Nielsen, **O Peralta**
  Parisian types of ruin probabilities for a class of dependent risk-reserve processes. Scandinavian Actuarial Journal. Advance online publication. doi:10.1080/03461238.2018.1483420

  This paper contains part of the results in Chapter 4 and Chapter 5.

- N Bean, G Nguyen, BF Nielsen, **O Peralta**
  Fluid RAP and queues. Manuscript in preparation.

  This paper contains part of the results in Chapter 7.

Chapter 2

Background Theory

In this chapter we develop the mathematical tools necessary to understand the results stemming from this thesis in further chapters. In Section 2.1 we give a brief overview of the theory of Markov processes. In particular, we study the Markov jump processes with countable state–space, and the piecewise deterministic Markov process introduced by Davis (1984). In Section 2.2 we review two classes of probability distributions whose study relies in matrix–analytic methods and probabilistic interpretations: the classes of phase–type (PH) and matrix–exponential (ME) distributions. In Section 2.3 we construct from scratch two classes of point processes related to the PH and ME distributions: the Markovian arrival process and the Rational arrival process, the former defined in Neuts (1979) and the latter in Asmussen and Bladt (1999). In Section 2.4 we review the so–called fluid flow process, whose rigorous study was initiated simultaneously by Asmussen (1995a), Rogers (1994) and Karandikar and Kulkarni (1995). Finally, in Section 2.5 we define several classes of risk process which have been used extensively in the literature: the Cramér–Lundberg process, the Sparre–Andersen process, spectrally negative Lévy processes, risk processes driven by a Markovian arrival process, and risk processes induced by a fluid flow process.

2.1 Markov processes

In this section we study a class of stochastic processes of crucial importance, not only in this thesis, but also within probability theory as a whole: the class of continuous–time and time–homogeneous Markov processes.

Let $(\mathfrak{E}, \mathcal{E})$ be a measurable space. Furthermore, let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0})$ be a filtered measurable space and $\{\mathbb{P}_x\}_{x\in\mathfrak{E}}$ a family of probability measures over it. Let $\mathrm{b}\mathcal{E}$ be the set of bounded $\mathcal{E}$-measurable functions, and for each $t \geq 0$ define the operator $P_t : \mathrm{b}\mathcal{E} \to \mathrm{b}\mathcal{E}$ by

$$(P_t f)(x) = \mathbb{E}_x(f(X_t)).$$

A **time–homogeneous Markov process** $\{X_t\}_{t\geq 0}$ with state–space $(\mathfrak{E}, \mathcal{E})$ is an $\mathfrak{E}$-valued $\mathcal{F}_t$–adapted stochastic process defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \{\mathbb{P}_x\}_{x\in\mathfrak{E}})$ such that

- $\mathbb{P}_x(X_0 = x) = 1$,
- For all $f \in \mathrm{b}\mathcal{E}$, $x \in \mathfrak{E}$ and $s, t \geq 0$,

  $$\mathbb{E}_x(f(X_{t+s}) \mid \mathcal{F}_t) = (P_s f)(X_t), \quad \mathbb{P}_x - \text{a.s.};$$

  this is called the **Markov property**.

In all the cases considered in this manuscript, $\mathfrak{E}$ will be a topological space and $\mathcal{E}$ its Borel σ–algebra. Furthermore, for all $t \geq 0$ we will take $\mathcal{F}_t = \sigma(\{X_s\}_{s\leq t})$ and $\mathcal{F} = \sigma(\{X_s\}_{s\geq 0})$. From now on we will omit "time–homogeneous" from "time–homogeneous Markov process", since all the Markov processes considered in this manuscript will be of this particular type.

In some instances it is convenient to use the Markov property "at random times", for example, at the time $\{X_t\}_{t\geq 0}$ exits or enters certain set for the first time. More generally, consider a random variable $T : \Omega \to [0, \infty)$ such that $\{T \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$: such a random variable is called a $\mathcal{F}_t$–**stopping time**. Now, for each $\mathcal{F}_t$–stopping time T define the σ–algebra

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0\}.$$

The process $\{X_t\}_{t\geq 0}$ is said to have the **strong Markov property** if for all $f \in \mathrm{b}\mathcal{E}$, $x \in \mathfrak{E}$, $s \geq 0$ and $\{\mathcal{F}_t\}$–stopping time T,

$$\mathbb{E}_x(f(X_{T+s}) \mid \mathcal{F}_T) = (P_s f)(X_T), \quad \mathbb{P}_x - \text{a.s. on } \{T < \infty\}.$$

The theory of general Markov processes is vast and complex on its own, see Dynkin (1965) for a classic study on the topic. In the following, we develop the theory of two specific types of Markov processes: the Markov jump process on a countable state–space and the piecewise deterministic Markov process. These two classes of Markov processes will be the building blocks of the theory in forthcoming sections and chapters of this manuscript.

2.1.1 Markov jump processes on a discrete state–space

Let $\{J_t\}_{t\geq 0}$ be a time–homogeneous càdlàg Markov process which takes values in some countable state–space $\mathfrak{E}$ endowed with the discrete topology. The process $\{J_t\}_{t\geq 0}$ is called a **Markov jump process**. The following is a pathwise description of such a process.

Define $T_0 = 0$ and let $T_1, T_2, \ldots$ denote the successive times at which $\{J_t\}_{t\geq 0}$ switches states. Define the discrete–time process $\{Y_n\}_{n\geq 0}$ by $Y_n = J_{T_n}$ for all $n \geq 0$. Then $\{Y_n\}_{n\geq 0}$ is the Markov chain of the succesive states visited by $\{J_t\}_{t\geq 0}$; if $\{J_t\}_{t\geq 0}$ ever gets absorbed in some state, say at the m-th jump, then define $Y_{m+1} = Y_{m+2} = \cdots = Y_m$. Let $\boldsymbol{Q} = \{q_{ij}\}_{i,j\in\mathfrak{E}}$ be the transition matrix of $\{Y_n\}_{n\geq 0}$.

The pathwise behaviour of $\{J_t\}_{t\geq 0}$ is characterised by the following (see Section 1.3 of Bladt and Nielsen (2017) for further details). The conditional distribution of $T_{n+1} - T_n$ given the event $\{Y_n = i\}$ is exponential with a certain parameter that depends on i, say $\lambda_i \geq 0$. This is equivalent to say that $\mathbb{P}(J_{t+\mathrm{d}t} \neq i | J_t = i) = \lambda_i \mathrm{d}t$. λ_i is called the **intensity of jump from the state** i. Now, given that there exists a jump in the interval $[t, t+\mathrm{d}t)$, it will land in $j \in \mathfrak{E}$ with probability q_{ij}: this means that $\mathbb{P}(J_{t+\mathrm{d}t} = j | J_t = i) = \lambda_i q_{ij} \mathrm{d}t$. Thus, for $i \neq j$ we may define $\lambda_{ij} := \lambda_i q_{ij}$ and call it the **intensity of jump from state** i **to state** j. See Figure 2.1 for a realization of a Markov jump process. Define $\lambda_{ii} = -\lambda_i$ and $\boldsymbol{\Lambda} = \{\lambda_{ij}\}_{i,j\in\mathfrak{E}}$. Then $\boldsymbol{\Lambda}$ is called the **intensity matrix of** $\{J_t\}_{t\geq 0}$. Notice that by definition $\boldsymbol{\Lambda}\boldsymbol{e} = \boldsymbol{0}$. If we let $p_{ij}^t := \mathbb{P}(J_t = j | J_0 = i)$ and $\boldsymbol{P}^t := \{p_{ij}^t\}_{i,j\in\mathfrak{E}}$, it can be shown (see Corollary 1.3.11 of Bladt and Nielsen (2017)) that under the condition $\sup_{i\in\mathfrak{E}}\{-\lambda_{ii}\} < \infty$,

$$\boldsymbol{P}^t = \exp(\boldsymbol{\Lambda} t).$$

In this case $\exp(\boldsymbol{A})$ denotes the exponential of the matrix $\boldsymbol{A}$, which may be defined by

$$\exp(\boldsymbol{A}) := \sum_{k=0}^{\infty} \frac{\boldsymbol{A}^k}{k!},$$

and is alternatively denoted by $e^{\boldsymbol{A}}$.

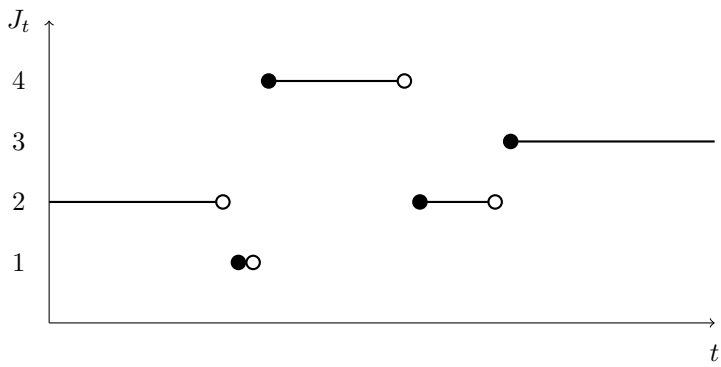


Figure 2.1: An example of a path of a Markov jump process with state space $\mathfrak{E} = \{1, 2, 3, 4\}$. Each holding time in $i \in \mathfrak{E}$ is exponentially distributed with parameter λ_i, and a jump from i to j occurs with probability $q_{ij} = \lambda_{ij}/\lambda_i$.

We classify the states of a Markov jump process $\{J_t\}_{t\geq 0}$ in accordance to the classification of those of $\{Y_n\}_{n\geq 0}$. More specifically, we say that a state $i \in \mathfrak{E}$ is transient (absorbing) for $\{J_t\}_{t\geq 0}$ if and only if i is transient (absorbing) for $\{Y_n\}_{n\geq 0}$. Notice that if some state $i \in \mathfrak{E}$ is absorbing, then $q_{ij} = \delta_{ij}$ and $\lambda_{ij} = 0$ for all $j \in \mathfrak{E}$.

Remark 1 *The strong Markov property holds for $\{J_t\}_{t\geq 0}$; see Asmussen (2003) for a direct proof of this statement. In this manuscript we take a longer path and prove it for a richer class of stochastic processes in Subsection 2.1.3, rendering the strong Markov property for Markov jump processes as a corollary.*

Markov jump processes represent a nice and tractable tool to construct complex stochastic systems. In particular, we will use them in the following sections of this chapter to construct the class of phase–type distributions, the Markovian arrival process and the fluid flow process.

2.1.2 The general jump process

In this subsection we briefly draw attention to a class of stochastic processes which are piecewise–constant and possibly non–Markovian. Such a class, called general jump processes, yields a great variety of models with significant complexity. In this subsection we follow part of the work of Davis (1976). We precisely define the general jump process and we study some properties of its natural

filtration and stopping times. The main result of this subsection is stated in Theorem 2.3, which provides a characterisation of all stopping times of the general jump process; we refrain from giving a full proof of such a characterisation, given that it goes well beyond the scope of this thesis and it adds no insight on any other topic treated here.

Let $\{X_t\}_{t\geq 0}$ be a right-continuous piecewise-constant stochastic process taking values in $\mathfrak{E}\cup\{\Delta_\infty\}$, where $\mathfrak{E}$ is a topological space endowed with its Borel σ–algebra $\mathcal{E}$, and Δ_∞ is an isolated state. Assume that $\{X_t\}_{t\geq 0}$ has discontinuities at a strictly increasing sequence of times $\{T_k\}_{k\geq 1}$ and is possibly killed at time $T_\infty := \lim_{k\to\infty} T_k$. The path of $\{X_t\}_{t\geq 0}$ can be specified by giving a fixed point $z_0 \in \mathfrak{E}$ and a possibly finite sequence $\{(S_k, Z_k)\}_{k\geq 1}$ of random vectors with $S_k \in \mathbb{R}_+$, $Z_k \in \mathfrak{E}$, defining $T_0 = 0$, $T_k = T_{k-1} + S_k$ and $X_t = Z_k$ for $t \in [T_k, T_{k+1})$ (with $T_{i+1} := \infty$ if the sequence $\{(S_k, Z_k)\}_{k\geq 1}$ has exactly i elements), and $X_t = \Delta_\infty$ for $t \geq T_\infty$. Assume that $\mathbb{P}(Z_{k-1} = Z_k) = 0$ for all k, so that "jumps are effectively jumps". The stochastic process $\{X\}_{t\geq 0}$ is called the **general jump process**. Notice that the Markov property is not necessarily attained for such a process.

Now, let us define the general jump process $\{X_t\}_{t\geq 0}$ in a canonical probability space. For $k \in \mathbb{N}$, let

$$\mathfrak{Y}_k = (\mathbb{R}_+ \times \mathfrak{E}) \cup \{\Delta\},$$

where Δ is an isolated point and let $\mathcal{Y}_k$ denote the Borel sets of $\mathfrak{Y}_k$. Define

$$\Omega_i = \prod_{k=1}^{i} \mathfrak{Y}_k, \quad \Omega = \prod_{k=1}^{\infty} \mathfrak{Y}_k, \quad \mathcal{F}^{i,0} = \bigotimes_{k=1}^{i} \mathcal{Y}_k \quad \textit{and} \quad \mathcal{F}^0 = \bigvee_{i=1}^{\infty} \mathcal{F}^{i,0},$$

where $\bigotimes_{k=1}^{i} \mathcal{Y}_k$ denotes the product σ–algebra of $\{\mathcal{Y}_k\}_{k=1}^{i}$, and $\bigvee_{i=1}^{\infty} \mathcal{F}^{i,0}$ is the smallest σ–algebra containing $\cup_{i=1}^{\infty} \mathcal{F}^{i,0}$.

Let $\xi_k : \Omega \to \mathfrak{Y}_k$ denote the k–th coordinate mapping. For $\omega \in \Omega$ such that $\xi_k(\omega) \in \mathbb{R}_+ \times \mathfrak{E}$ define $(S_k(\omega), Z_k(\omega)) := \xi_k(\omega)$. For any $\omega \in \Omega$, let $\omega_k(\omega) = (\xi_1(\omega), \ldots, \xi_k(\omega))$, let

$$T_k(\omega) := \begin{cases} \sum_{i=1}^{k} S_i(\omega), & \xi_i(\omega) \neq \Delta, i = 1, \ldots, k \\ \infty, & \xi_i(\omega) = \Delta \text{ for some } i = 1, \ldots, k, \end{cases} \quad \text{and}$$

$$T_\infty(\omega) := \lim_{k\to\infty} T_k(\omega),$$

and for $t \in \mathbb{R}_+$ define

$$X_t(\omega) = \begin{cases} z_0, & t < T_1(\omega) \\ Z_k, & T_k(\omega) \leq t < T_{k+1}(\omega) \\ \Delta_\infty, & t \geq T_\infty(\omega), \end{cases}$$

which determines the sample path of $\{X_t\}_{t\geq 0}$.

Let $\{\mathcal{F}_t^0\}_{t\geq 0}$ be such that $\mathcal{F}_t^0 = \sigma(\{X_s\}_{s\leq t})$. We characterise the probability measure $\mathbb{P}$ on Ω by giving the following family of conditional distributions: μ^1 is a measure over $(\mathfrak{Y}_1, \mathcal{Y}_1)$ such that

$$\mu^1((\{0\} \times \mathfrak{E}) \cup (\mathbb{R}_+ \times \{z_0\})) = 0.$$

For $k \geq 2$, $\mu^k : \Omega_{k-1} \times \mathcal{Y}_k \to [0,1]$ is a transition measure satisfying:

1. $\mu^k(\cdot; A)$ is measurable for each $A \in \mathcal{Y}_k$;
2. $\mu^k(\omega_{k-1}(\omega); \cdot)$ is a probability measure for each $\omega \in \Omega$;
3. $\mu^k(\omega_{k-1}(\omega); (\{0\} \times \mathfrak{E})) = 0$ for each $\omega \in \Omega$, so that no jumps occur at time 0;
4. $\mu^k(\omega_{k-1}(\omega); (\mathbb{R}_+ \times Z_{k-1}(\omega))) = 0$ for each $\omega \in \Omega$, so that once a jump happens from some point $a \in \mathfrak{E}$, it lands in somewhere different to a;
5. $\mu^k(\omega_{k-1}(\omega); \{\Delta\}) = 1$ if $\xi_i(\omega) = \Delta$ for some $i \leq k-1$, so that once the process reaches Δ, it gets absorbed there.

We define $\mathbb{P}$ to be the unique probability measure on $(\Omega, \mathcal{F}^0)$ such that for each $k \geq 1$ and bounded measurable function f over $(\Omega_k, \mathcal{F}^{k,0})$,

$$\begin{aligned}&\int_\omega f(\xi_1(\omega), \ldots, \xi_k(\omega))\mathbb{P}(\mathrm{d}\omega)\\ &\quad = \int_{\mathfrak{Y}_1} \cdots \int_{\mathfrak{Y}_k} f(\xi_1, \ldots, \xi_k)\mu^k(\xi_1, \ldots, \xi_{k-1}; \mathrm{d}\xi_k) \cdots \mu^1(\mathrm{d}\xi_1).\end{aligned}$$

Let $\mathcal{F}_t$, $\mathcal{F}^k$ and $\mathcal{F}$ be the σ-algebras $\mathcal{F}_t^0$, $\mathcal{F}^{k,0}$ and $\mathcal{F}^0$ completed with the $\mathbb{P}$-null sets of $\mathcal{F}^0$, respectively.

Lemma 2.1 *1. T_k, $k \in \mathbb{N}$, and T_∞ are $\mathcal{F}_t$-stopping times;*
2. $\mathcal{F} = \mathcal{F}_\infty$, where $\mathcal{F}_\infty := \vee_{t\geq 0}\mathcal{F}_t$.

Proof.

1. With probability 1 we can say if the k-th jump has occured by time t or not by inspecting $\{X_s\}_{s\leq t}$; the completion of $\mathcal{F}_t^0$ into $\mathcal{F}_t$ yields the desired result.
2. That $\mathcal{F}_\infty \subset \mathcal{F}$ is immediate from the definition. The converse follows by noting that ξ_i is $\mathcal{F}_\infty$-measurable for all i, which in turn follows from $\{T_i\}_{i\geq 1}$ being $\mathcal{F}_t$-stopping times.

□

For any $\mathcal{F}_t$-stopping time, let

$$\mathcal{F}_T := \{A \in \mathcal{F} : A \cap (T \leq t) \in \mathcal{F}_t \text{ for all } t \geq 0\}.$$

Then we have the following.

Theorem 2.2 *1.* $\mathcal{F}_t = \cap_{\epsilon>0}\mathcal{F}_{t+\epsilon}$.

2. For any $\mathcal{F}_t$*–stopping time* T, $\mathcal{F}_T = \sigma(X_{s\wedge T}, s \geq 0)$.

3. For each $k \geq 1$, $\mathcal{F}_{T_k} = \sigma(\xi_1, \ldots, \xi_k)$.

Proof. 1) and 2) are valid for every right-constant process. 3) follows by using 2) and the 1–to–1 correspondence between $(\xi_1, \ldots, \xi_k)$ and $(X_{t\wedge T_k}, t \geq 0)$. □

The following is a characterisation of all the $\mathcal{F}_t$–stopping times of $\{X_t\}_{t\geq 0}$; its proof can be found in Davis (1993) (Theorem A2.3).

Theorem 2.3 *Let* T *be a* $\mathcal{F}_t$*-stopping time. Then there exists a constant* s_1 *and* $\mathcal{F}^k$*–measurable functions* $s_k : \Omega_{k-1} \to \mathbb{R}_+$ *(*$k \geq 2$*) such that*

$$T\mathbb{1}_{\{T\leq T_1\}} = (s_1 \wedge T_1)\mathbb{1}\{T \leq T_1\},$$

and for $k \geq 2$

$$T\mathbb{1}_{\{T_{k-1}<T\leq T_k\}} = ((T_{k-1} + s_k(\xi_{k-1})) \wedge T_k)\mathbb{1}\{T_{k-1} < T \leq T_k\}.$$

Theorem 2.3 provides an explicit characterisation of **any** $\mathcal{F}_t$–stopping time T associated to a general jump process. In the following section, this will aid us in proving the strong Markov property for the class of piecewise deterministic Markov processes.

2.1.3 Piecewise deterministic Markov processes

In Davis (1984), a general class of non–diffusion Markovian models is formulated, called piecewise deterministic Markov processes. In this subsection we study its construction and prove that under mild conditions, such a class has the strong Markov property. We follow closely the classic text of Davis (1993); Jacobsen (2006) and Rudnicki and Tyran-Kamińska (2017) are alternative references to the subject, the latter being oriented to applications. Piecewise deterministic

Markov processes and their properties will play a central role in the construction of the Rational arrival process and the Fluid RAP (Chapter 7).

Let $g : \mathbb{R}^d \to \mathbb{R}^d$ be a Lipschitz continuous function. For such a function, the ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = g(x(t)), \quad x(0) = x \in \mathbb{R}^d \tag{2.1.1}$$

has a unique solution defined for all $t \in \mathbb{R}$ (see Chapter 17 of Hirsch et al. (2013)). Denote this solution by $\phi(t, x)$, where the second entry emphasizes the dependence on the initial point x. Since $\phi_t^{-1}(x) = \phi(-t, x)$ for any $x \in \mathbb{R}^d$, then the map $\phi_t : x \to \phi(t, x)$ is bijective. Furthermore, the family $\{\phi_t\}_{t\in\mathbb{R}}$ has the semigroup property, that is, $\phi_{t+s} = \phi_t \circ \phi_s$ for all $s, t \geq 0$.

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a C^1 function and let $x(t) = \phi(t, x)$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}f(x(t)) = \sum_i \frac{\partial f}{\partial x^i}(x(t))g^i(x(t)).$$

Let $\mathfrak{X}$ denote the first-order differential operator

$$\mathfrak{X}f(x) = \sum_i \frac{\partial f}{\partial x^i}(x)g^i(x).$$

Then $x(t)$ satisfies (2.1.1) if and only if it satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}f(x(t)) = \mathfrak{X}f(x(t)) \quad \text{for all} \quad f \in C^\infty(\mathbb{R}^d). \tag{2.1.2}$$

We say that (2.1.2) is the **coordinate free** form of the differential equation, $\mathfrak{X}$ its **vector field** and $\phi(t, x)$ is the **flow** of $\mathfrak{X}$.

Now suppose that g is only locally Lipschitz continuous, that is, if over any compact set $D \subset \mathbb{R}^d$ there exists a constant K_D such that

$$|g(x) - g(y)| \leq K_D|x - y|, \quad x, y \in D.$$

Under this condition, the solution $x(t)$ to (2.1.1) starting at $x \in D$ is uniquely determined for $t \leq t_D(x)$, where $t_D(x)$ is the time at which the solution exits from D. For $x \in \mathbb{R}^d$, define

$$t_\infty(x) := \lim_{r\to\infty} t_{D_r}(x),$$

where $D_r = \{x \in \mathbb{R}^d : \|x\|_{\max} \leq r\}$. An **explosion** is said to happen at $t_\infty(x) < \infty$ if $|x(t)| \to \infty$ as $t \uparrow t_\infty$. From now on, assume that the vector fields

we consider are locally Lipschitz continuous and that for every initial point $x \in \mathbb{R}^d$ no explosion occurs.

Let $K \subseteq \mathbb{N}$. For each $\nu \in K$, let $\mathfrak{E}_\nu^0$ be an open subset of $\mathbb{R}^\nu$ and define

$$\mathfrak{E}^0 := \cup_{\nu \in K} \mathfrak{E}_\nu^0 = \{(\nu, \zeta) : \nu \in K, \zeta \in \mathfrak{E}_\nu^0\}.$$

Notice that by the way $\mathfrak{E}^0$ is defined, it is a disjoint union in a set–theoretic sense; in a geometric sense, $\mathfrak{E}^0$ is the union of subsets of Euclidian spaces of varying dimension. For each $\nu \in K$, let $\mathfrak{X}_\nu$ be a locally Lipschitz continuous vector field in $\mathfrak{E}_\nu^0$ which in turn determines a flow $\phi_\nu(t, \zeta)$. From now on, any function $h : \mathfrak{E}^0 \to \mathbb{R}$ is identified with its component functions $h_\nu : \mathfrak{E}_\nu^0 \to \mathbb{R}$ and we write $\mathfrak{X}h(x)$ in place of $\mathfrak{X}_\nu h_\nu(\zeta)$ for $x = (\nu, \zeta) \in \mathfrak{E}^0$. For such an x, define

$$t_*(x) := \inf\{t > 0 : \phi_\nu(t, \zeta) \in \partial\mathfrak{E}_\nu^0\},$$

where $\partial\mathfrak{E}_\nu^0$ denotes the boundary of $\mathfrak{E}_\nu^0$. Since t_* is the hitting time to a closed set, it is measurable (see Lemma 27.1 in Davis (1993)). To rule out explosions, assume that $t_\infty(x) = \infty$ whenever $t_*(x) = \infty$. Define

$$\partial^\pm \mathfrak{E}_\nu^0 = \{z \in \partial\mathfrak{E}_\nu^0 : z = \phi_\nu(\pm t, \zeta) \text{ for some } \zeta \in \mathfrak{E}_\nu^0, t > 0\}$$

and $\partial_1 \mathfrak{E}_\nu^0 = \partial^- \mathfrak{E}_\nu^0 \setminus \partial^+ \mathfrak{E}_\nu^0$. Let $\mathfrak{E}_\nu = \mathfrak{E}_\nu^0 \cup \partial_1 \mathfrak{E}_\nu^0$, $\mathfrak{E} = \cup_\nu \mathfrak{E}_\nu$ and

$$\Gamma^* = \cup_\nu \partial^+ \mathfrak{E}_\nu^0.$$

We endow the set $\mathfrak{E}$ with the σ-algebra $\mathcal{E}$ composed of subsets $A \subseteq \mathfrak{E}$ taking the form $A = \cup_\nu A_\nu$, where A_ν is a Borel set of $\mathfrak{E}_\nu$.

The **piecewise deterministic Markov process (PDMP)** with state–space $\mathfrak{E}$ is a stochastic process $\{X_t\}_{t \geq 0}$ which is governed by random jumps at points in time, and between those points its evolution is deterministically governed the vector field $\mathfrak{X}$. The jump mechanism is determined by two functions, the **jump rate** λ and the **transition measure** Q. $(\mathfrak{X}, \lambda, Q)$ are called the **local caracteristics** of the PDMP. The jump rate $\lambda : \mathfrak{E} \to \mathbb{R}_+$ is a measurable function which for each $x = (\nu, \zeta) \in \mathfrak{E}$ there exists $\epsilon(x) > 0$ such that the function $s \to \lambda(\nu, \phi_\nu(s, \zeta))$ is integrable on $[0, \epsilon(x))$. The transition measure $Q : \mathfrak{E} \times \mathcal{E} \cup \Gamma^* \to [0, 1]$ is such that

1. $Q(\cdot; A)$ is measurable for each fixed A,
2. $Q(x; \cdot)$ is a probability measure for each $x \in \mathcal{E} \cup \Gamma^*$, and
3. $Q(x; \{x\}) = 0$.

For $f \in L_1(Q(x; \cdot))$, define

$$Qf(x) = \int_{\mathfrak{E}} f(y) Q(x; \mathrm{d}y).$$

Notice that as an operator, Q maps set of bounded measurable functions on $\mathfrak{E} \cup \Gamma^*$ into itself.

In order to properly define the PDMP $\{X_t\}_{t\geq 0}$, let $(\Omega, \mathcal{A}, \mathbb{P})$ be the Hilbert cube, that is, the canonical space for a sequence $\{U_i\}_{i\geq 1}$ of independent $U(0,1)$-distributed random variables. For each $\omega \in \Omega$, the sample path $\{X_t(\omega)\}_{t\geq 0}$ with values in $\mathfrak{E}$, starting from a fixed point $x = (\nu, \zeta) \in \mathfrak{E}$ is defined as follows. Let

$$F(t,x) = \mathbb{1}_{t<t_*(x)} \exp\left(-\int_0^t \lambda(\nu, \phi_\nu(s,\zeta))\mathrm{d}s\right) \tag{2.1.3}$$

be the survivor function of the first jump time T_1. Let

$$\psi_1(u,x) = \inf\{t : F(t,x) \leq u\},$$

and for $\omega \in \Omega$ define $S_1(\omega) = T_1(\omega) = \psi_1(U_1(\omega), x)$. Let $\psi_2 : [0,1]\times(\mathfrak{E}\cup\Gamma^*) \to \mathfrak{E}$ be a measurable function such that $l\{u : \psi_2(u,x) \in A\} = Q(x;A)$ for $A \in \mathcal{E}$, where l denotes the Lebesgue measure on $[0,1]$; see Corollary 23.4 in Davis (1993) for the proof of existance of such a function. For all $\omega \in \Omega$ we define

$$X_t(\omega) = \begin{cases} (\nu, \phi_\nu(t,\zeta)), & t \geq 0 \quad \text{if} \quad T_1(\omega) = \infty \\ (\nu, \phi_\nu(t,\zeta)), & 0 \leq t < T_1(\omega) \quad \text{if} \quad T_1(\omega) < \infty, \quad \text{and} \end{cases}$$
$$X_{T_1}(\omega) = \psi_2(U_2(\omega), (\nu, \phi_\nu(S_1(\omega), \zeta))),$$

leaving the process $\{X_t(\omega)\}_{t\geq 0}$ defined up to the first jump time $T_1(\omega)$. If $T_1(\omega) < \infty$, the process restarts from $X_{T_1}(\omega)$ according to the same recipe. More specifically, define

$$S_2(\omega) = \psi_1(U_3(\omega), X_{T_1}(\omega)), \quad T_2(\omega) = T_1(\omega) + S_2(\omega), \quad \text{and} \quad (\nu', \zeta') = X_{T_1}(\omega),$$

and let

$$X_t(\omega) = \begin{cases} (\nu', \phi_{\nu'}(t - T_1(\omega), \zeta')), & t \geq T_1 \quad \text{if} \quad T_2(\omega) = \infty \\ (\nu', \phi_{\nu'}(t - T_1(\omega), \zeta')), & T_1 \leq t < T_2, \quad \text{if} \quad T_2(\omega) < \infty, \quad \text{and} \end{cases}$$
$$X_{T_2}(\omega) = \psi_2(U_4(\omega), (\nu', \phi_{\nu'}(S_2(\omega), \zeta'))).$$

This will leave the process $\{X_t(\omega)\}_{t\geq 0}$ defined up to the second jump time $T_2(\omega)$. Repeating this procedure, we will produce two possibly finite sequences $(S_1(\omega), S_2(\omega), \ldots, S_{k_0}(\omega))$ and $(Z_1(\omega), Z_2(\omega), \ldots, Z_{k_0-1}(\omega))$ where $k_0(\omega) = \min\{k : S_k(\omega) = \infty\}$ and $Z_k(\omega) = X_{T_k}(\omega)$.

Define

$$N_t = \sum_k \mathbb{1}_{t\geq T_k},$$

the number of jumps by time t. In this manuscript we will assume that for every starting point $x \in \mathfrak{E}$ and $t \geq 0$, $\mathbb{E}(N_t) < \infty$. If this happens, then there will not be infinite jump activity in any compact interval, and thus X_t will be well–defined for all $t \geq 0$. The next condition ensures that $\mathbb{E}(N_t) < \infty$.

Proposition 2.4 *Suppose that $\lambda(x) \leq c < \infty$ for all $x \in \mathfrak{E}$. If there are no active boundaries (that is, $t_*(x) = \infty$ for all $x \in \mathfrak{E}$) then $\mathbb{E}(N_t) < \infty$.*

PROOF. Majorizing with a Poisson process of rate c we have that $\mathbb{E}(N_t) \leq ct$ and the proof is finished. □

From now on, suppose that $\mathbb{E}(N_t) < \infty$ for all $t \geq 0$.

The construction of $\{X_t\}_{t\geq 0}$ defines for each $x \in \mathfrak{E}$ a measurable mapping $\gamma_x : \Omega \to \mathscr{D}_{\mathfrak{E}}$, where $\mathscr{D}_{\mathfrak{E}}$ is the set of cádlág functions which take values in $\mathfrak{E}$. Let $\mathbb{P}_x := \mathbb{P}\gamma_x^{-1}$: this defines a family of measures $\{\mathbb{P}_x\}_{x\in\mathfrak{E}}$ over $\mathscr{D}_{\mathfrak{E}}$. Thus, a PDMP can be thought either as a process defined on Ω or $\mathscr{D}_{\mathfrak{E}}$. Now, for any probability measure μ on $\mathfrak{E}$, define a measure on $(\mathscr{D}_{\mathfrak{E}}, \mathcal{F}^0)$ (with $\mathcal{F}_t^0 = \sigma(X_s, s \leq t)$ and $\mathcal{F}^0 = \vee_t \mathcal{F}_t^0$) by

$$\mathbb{P}^\mu(A) = \int_{\mathfrak{E}} \mathbb{P}_x(A)\mu(\mathrm{d}x).$$

Let $\mathcal{F}_t^\mu$ be the completion of $\mathcal{F}_t^0$ with all $\mathbb{P}^\mu$-null sets of $\mathcal{F}^0$, and define

$$\mathcal{F}_t = \cap_{\mu\in\mathscr{P}(\mathfrak{E})}\mathcal{F}_t^\mu,$$

where $\mathscr{P}(\mathfrak{E})$ denotes the set of probability measures over $(\mathfrak{E}, \mathcal{E})$.

Now, let us construct the so–called **associated general jump process** of the PDMP started at $x \in \mathfrak{E}$. Such a process, denoted by $\{\eta_t\}_{t\geq 0}$, takes values in $\mathfrak{E} \times \mathbb{Z}_+$ and is defined by

$$\eta_t = \begin{pmatrix} x \\ 0 \end{pmatrix}, \quad \text{for } t < T_1,$$
$$\eta_t = \begin{pmatrix} \eta_t^1 \\ \eta_t^2 \end{pmatrix} := \begin{pmatrix} X_{T_n} \\ n \end{pmatrix}, \quad \text{for } T_n \leq t < T_{n+1}.$$

Notice that there is a 1–to–1 correspondence between the sample paths of $\{X_t\}_{t\geq 0}$ and $\{\eta_t\}_{t\geq 0}$.

Theorem 2.5 *$\{\mathcal{F}_t\}_{t\geq 0}$ is right continuous.*

PROOF. Follows from Theorem 2.2. □

For bounded measurable functions $f : \mathfrak{E} \to \mathbb{R}$, denote

$$P_t f(x) = \mathbb{E}_x(f(X_t)).$$

Theorem 2.6 *The process $\{X_t\}_{t\geq 0}$ is a time–homogeneous strong Markov process, that is, for any $x \in \mathfrak{E}$, $\mathcal{F}_t$-stopping time T and bounded measurable function f,*

$$\mathbb{E}_x(f(X_{T+s})\mathbb{1}_{T<\infty}|\mathcal{F}_T) = P_s f(X_T)\mathbb{1}_{T<\infty}. \tag{2.1.4}$$

Proof. We will first show that the simple Markov property holds, that is, that (2.1.4) holds with T replaced by a fixed $t > 0$. By construction and by Theorem 2.2, we have that on the event $\{T_k < \infty\}$, $\mathbb{P}(S_{k+1} > s|\mathcal{F}_{T_k}) = F(s, X_{T_k})$. Then,

$$\begin{aligned}\mathbb{P}_x(T_{k+1} > t+s|\mathcal{F}_t)\mathbb{1}_{T_k\leq t<T_{k+1}} &= \frac{F(t+s-T_k, X_{T_k})}{F(t-T_k, X_{T_k})}\mathbb{1}_{T_k\leq t<T_{k+1}} \\ &= \exp\left(-\int_0^s \lambda(\nu, \phi_\nu(u+t-T_k, \zeta))\mathrm{d}u\right)\mathbb{1}_{t+s-T_k<t_*(X_{T_k})}\mathbb{1}_{T_k\leq t<T_{k+1}},\end{aligned}$$

where $(\nu, \zeta) = X_{T_k}$. Denote $(\nu(t), \zeta(t)) = X_t$. Then, on the event $\{T_k \leq t < T_{k+1}\}$ and $u > 0$,

$$\phi_\nu(u+t-T_k, \zeta) = \phi_\nu(u, \phi_\nu(t-T_k, \zeta)) = \phi_{\nu(t)}(u, \zeta(t)), \text{ and}$$

$$t_*(X_t) = t_*(X_{T_k}) - (t-T_k),$$

so that $(t+s-T_k) < t_*(X_{T_k})$ is equivalent to $s < t_*(X_t)$. Thus,

$$\mathbb{P}_x(T_{k+1} > t+s|\mathfrak{F}_t)\mathbb{1}_{T_k\leq t<T_{k+1}} = F(s, X_t)\mathbb{1}_{T_k\leq t<T_{k+1}}.$$

If we let $\hat{T} := \inf\{s > t : X_s \neq X_{s^-}\}$, we have that $\mathbb{P}_x(\hat{T} > t+s|\mathcal{F}_t) = F(s, X_t)$. Thus, conditional on $\mathcal{F}_t$, the next jump time of the process has the same distribution as the first jump time of a PDMP started at X_t, proving that $\{X_t\}_{t\geq 0}$ is a homogeneous Markov process, and thus (2.1.4) holds with $T = t$.

To prove the strong Markov property, use the 1–to–1 correspondence between $\{X_t\}_{t\geq 0}$ and $\{\eta_t\}_{t\geq 0}$, and use than any $\mathcal{F}_t$–stopping time associated to $\{\eta_t\}_{t\geq 0}$ can be decomposed as in Theorem 2.3. That is, for the stopping time T there exist measurable functions $s_1, s_2, s_3, \ldots$ such that

$$T = (s_1(x) \wedge T_1)\mathbb{1}_{T\leq T_1} + \sum_{k=2}^{\infty}((T_{k-1} + s_k(\xi_{k-1})) \wedge T_k)\mathbb{1}_{T_{k-1}<T\leq T_k}.$$

Thus, on the set $\{T < \infty\}$, there are three possible cases: either $T = 0$, or $T = T_k$ for some k, or $T = T_{k-1} + s_k(\xi_{k-1})$ for some k, and each one of these sets is $\mathcal{F}_T$-measurable. If we define $T^* = \inf\{t > T : X_t \neq X_{t^-}\}$, then we have that for any of those three cases $\mathbb{P}_x(T^* > T+s, T < \infty|\mathcal{F}_T) = F(s, X_T)\mathbb{1}_{T<\infty}$, thus the strong Markov property holds. That (2.1.4) holds is proved by a standard approximation of f by simple functions. □

2.2 Probability distributions

In this section we review some distributions over $\mathbb{R}_+$ which are suitable to study within a matrix–analytic framework: phase–type distributions and matrix–exponential distributions. We develop some key results and lay out the probabilistic approach used in forthcoming sections and chapters. Additionally, we study a class of multivariate phase–type distributions, that is, multivariate distributions which have phase–type–distributed marginals.

2.2.1 Phase–type distributions

One of the first attempts to generalize the exponential distribution was developed in Erlang (1909), where a distribution based on sequential i.i.d. hidden phases with exponential holding times was used to examine the number of telephone calls; such a distribution went on to be known as the Erlang distribution. Later on, the idea of hidden stages was generalized in Jensen (1954) to allow non–sequential hidden phases with possibly different intensities; this class of distributions went on to be named phase–type distributions. The theory of phase–type distributions had a significant growth in the seventies, mainly to the work of M. F. Neuts and co-authors, who laid out the basis of matrix–analytic methods in Applied Probability; a large portion of their work is summarized in Neuts (1994). Nowadays, research on phase–type distributions and its applications keeps flourishing, with prolific results in areas such as queueing theory (Asmussen (2003)) and risk theory (Asmussen and Albrecher (2010)). In the following we develop some basic results concerning this class of distributions.

Let $\{J_t\}_{t\geq 0}$ be a Markov jump process with state–space given by $\mathfrak{E} = \{1, \ldots, p, p+1\}$ such that the states $\{1, \ldots, p\}$ are transitory and $p+1$ is absorbing. This means that the intensity matrix associated to $\{J_t\}_{t\geq 0}$ is on the form

$$\begin{pmatrix} \boldsymbol{T} & \boldsymbol{t} \\ \boldsymbol{0} & 0 \end{pmatrix},$$

with $\boldsymbol{t} = -\boldsymbol{T}\boldsymbol{e}$. Furthermore, let $J_0 \in \{1, \ldots, p\}$ be chosen according to some probability vector distribution $\boldsymbol{\pi}$, which is expressed in a row–vector fashion. The matrix $\boldsymbol{T} = \{t_{ij}\}_{i,j\in\{1,\ldots,p\}}$ will be called the **sub–intenisty matrix** and $\boldsymbol{\pi} = \{\pi_i\}_{i\in\{1,\ldots,p\}}$ the **initial distribution vector**. Additionally, $\boldsymbol{t} = \{t_i\}_{i\in\{1,\ldots,p\}}$ will be called the **absorption intensities vector** . Notice that the pair $(\boldsymbol{\pi}, \boldsymbol{T})$ characterise the distributional properties of $\{J_t\}_{t\geq 0}$.

Let

$$\tau = \inf\{t > 0 : J_t = p + 1\},$$

that is, τ corresponds to the absorption time of $\{J_t\}_{t\geq 0}$. We say that τ follows a **phase–type distribution of parameters** $(\boldsymbol{\pi}, \boldsymbol{T})$ and we will denote it by $\tau \sim \mathrm{PH}_p(\boldsymbol{\pi}, \boldsymbol{T})$ or $\tau \sim \mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$, depending if we want to be explicit about the dimension of the representation $(\boldsymbol{\pi}, \boldsymbol{T})$ or not. We call the set of transient states of $\{J_t\}_{t\geq 0}$ the **phase–space** of $\mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$. See Figure 2.2 for an example of a realization of a phase–type–distributed random variable.

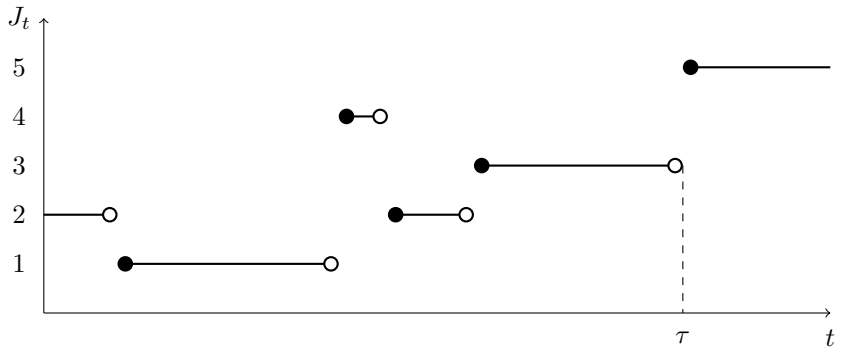


Figure 2.2: A realization of the underlying Markov jump process of the phase–type–distributed random variable τ with phase–space $\{1, 2, 3, 4\}$ and absorbing state 5. Notice that τ corresponds to the length of time it takes $\{J_t\}_{t\geq 0}$ to be absorbed in 5.

The following are some basic results concerning phase–type distributions which show the machinery and techniques used in a large portion of the proofs of this manuscript.

Theorem 2.7 *Let* $\tau \sim \mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$ *and* $x \geq 0$. *Denote by* $f_\tau(\cdot)$ *the density function of* τ. *Then*

1. $\mathbb{P}(\tau > x) = \boldsymbol{\pi} e^{\boldsymbol{T}x} \boldsymbol{e}$, *and*
2. $f_\tau(x) = \boldsymbol{\pi} e^{\boldsymbol{T}x} \boldsymbol{t}$.

Proof.

1. The standard series expansion of the matrix–exponential function reveals that
$$\exp\left(\begin{pmatrix} \boldsymbol{T} & \boldsymbol{t} \\ \boldsymbol{0} & 0 \end{pmatrix} x\right) = \begin{pmatrix} \exp(\boldsymbol{T}x) & \boldsymbol{e} - \exp(\boldsymbol{T}x)\boldsymbol{e} \\ \boldsymbol{0} & 1 \end{pmatrix}.$$

This implies that for $i, j \in \{1, \ldots, p\}$, $(\exp(\boldsymbol{T}x))_{ij}$ corresponds to the probability that $J_x = j$ given that $J_0 = i$. Then,

$$\begin{aligned}\mathbb{P}(\tau > x) &= \sum_{i=1}^{p}\sum_{j=1}^{p}\mathbb{P}(J_0 = i)\mathbb{P}(J_x = j \mid J_0 = i) \\ &= \sum_{i=1}^{p}\sum_{j=1}^{p}\pi_i \exp(\boldsymbol{T}x)_{ij} \\ &= \boldsymbol{\pi} e^{\boldsymbol{T}x}\boldsymbol{e}.\end{aligned}$$

2. We have that

$$\begin{aligned}f_\tau(x)\mathrm{d}x &= \mathbb{P}(\tau \in (x, x+\mathrm{d}x)) \\ &= \sum_{j=1}^{p}\mathbb{P}(J_x = j)\mathbb{P}(\tau \in (x, x+\mathrm{d}x) \mid J_x = j) \\ &= \sum_{j=1}^{p}\left(\boldsymbol{\pi}e^{\boldsymbol{T}x}\right)_j (t_j \mathrm{d}x) \\ &= \boldsymbol{\pi}e^{\boldsymbol{T}x}\boldsymbol{t}\mathrm{d}x,\end{aligned}$$

and the proof is finished.

□

Remark 2 *Sometimes it will be convenient to work with phase–type representations* $(\boldsymbol{\pi}, \boldsymbol{T})$ *such that* $\boldsymbol{\pi}\boldsymbol{e} \in (0, 1)$. *Such representations will be called* **defective**, *in the sense that their distribution have an "unassigned" point mass of size* $1 - \boldsymbol{\pi}\boldsymbol{e}$. *In most cases, this point mass is assigned to either* 0 *or* ∞.

The following is a slightly more technical proof, needed in the present section and in further chapters.

Theorem 2.8 *Let* $\{\boldsymbol{h}_i\}_{i=1}^{p}$ *be p-dimensional linearly independent row–vectors and* $\boldsymbol{A}$ *a* $p \times p$ *real matrix such that* $\lim_{x\to\infty}\boldsymbol{h}_i e^{\boldsymbol{A}x} = \boldsymbol{0}$. *Then,*

1. $\lim_{x\to\infty} e^{\boldsymbol{A}x} = \boldsymbol{0}$, *and*
2. *the* n *eigenvalues of* $\boldsymbol{A}$, *say* $\{\theta_i\}_{i=1}^{n}$, *are such that* $\mathrm{Re}(\theta_i) < 0$ *for all* $i \in \{1, \ldots, n\}$.

PROOF.

1. Let
$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{h}_1 \\ \vdots \\ \boldsymbol{h}_p \end{pmatrix}.$$
The inverse matrix of $\boldsymbol{H}$ exists, so that
$$\begin{aligned} \lim_{x\to\infty} e^{\boldsymbol{A}x} &= \boldsymbol{H}^{-1} \lim_{x\to\infty} \boldsymbol{H} e^{\boldsymbol{A}x} \\ &= \boldsymbol{H}^{-1} \lim_{x\to\infty} \begin{pmatrix} \boldsymbol{h}_1 e^{\boldsymbol{A}x} \\ \vdots \\ \boldsymbol{h}_p e^{\boldsymbol{A}x} \end{pmatrix} \\ &= \boldsymbol{H}^{-1}\boldsymbol{0} = \boldsymbol{0}. \end{aligned}$$
2. Let us prove it by contradiction. Suppose that $\boldsymbol{A}$ has an eigenvalue $\theta = \alpha + i\beta$ with $\alpha \geq 0$ and left eigenvector $\boldsymbol{v} = \boldsymbol{u} + i\boldsymbol{w}$, where either $\boldsymbol{u} \neq \boldsymbol{0}$ or $\boldsymbol{w} \neq \boldsymbol{0}$. Now,
$$\begin{aligned} \boldsymbol{v}e^{\boldsymbol{A}x} &= e^{\alpha x}(\cos(\beta x) + i\sin(\beta x))\boldsymbol{v}_1 \\ &= e^{\alpha x}([\cos(\beta x)\boldsymbol{u} - \sin(\beta x)\boldsymbol{w}] + i[\cos(\beta x)\boldsymbol{w} + \sin(\beta x)\boldsymbol{u}]). \end{aligned}$$
W.l.o.g. suppose that $\boldsymbol{u} \neq \boldsymbol{0}$. Then
$$\boldsymbol{u}e^{\boldsymbol{A}x} = e^{\alpha x}[\cos(\beta x)\boldsymbol{u} - \sin(\beta x)\boldsymbol{w}]. \tag{2.2.1}$$
Item 1) implies that the l.h.s. of (2.2.1) goes to 0 as $x \to \infty$, while the r.h.s. does not, which is a contradiction.

□

Theorem 2.9 *Let $\tau \sim \mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$. Then,*

1. *$\boldsymbol{T}$ is invertible,*
2. *$\mathbb{E}(\tau) = \boldsymbol{\pi}(-\boldsymbol{T})^{-1}\boldsymbol{e}$.*

PROOF.

1. Notice that transitivity of the associated Markov jump process implies that necessarily $\lim_{x\to\infty} \boldsymbol{e}_i' e^{\boldsymbol{T}x} = \boldsymbol{0}$. Thus, apply Theorem 2.8 with $\boldsymbol{h}_i$ replaced with $\boldsymbol{e}_i'$, and $\boldsymbol{A}$ replaced with $\boldsymbol{T}$.

2. For $i, j \in \{1, \ldots, p\}$, let u_{ij} be the expected time the underlying Markov jump process $\{J_t\}_{t \geq 0}$ spends in state j before apsorption given $J_0 = i$. Then

$$\begin{aligned} u_{ij} &= \mathbb{E}\left(\int_0^\tau \mathbb{1}\{J_x = j\} \mathrm{d}x \mid J_0 = i \right) \\ &= \int_0^\infty \mathbb{E}\left(\mathbb{1}\{J_x = j\} \mid J_0 = i \right) \mathrm{d}x \\ &= \int_0^\infty (e^{\boldsymbol{T}x})_{ij} \mathrm{d}x. \end{aligned}$$

If $\boldsymbol{U} = \{u_{ij}\}_{i,j \in \{1,\ldots,p\}}$, then

$$\boldsymbol{U} = \int_0^\infty e^{\boldsymbol{T}x} \mathrm{d}x = [\boldsymbol{T}^{-1} e^{\boldsymbol{T}x}]_0^\infty = -\boldsymbol{T}^{-1}.$$

Thus,

$$\mathbb{E}(\tau) = \sum_{i=1}^{p} \sum_{j=1}^{p} \pi_i u_{ij} = \boldsymbol{\pi} \boldsymbol{U} \boldsymbol{e} = \boldsymbol{\pi} (-\boldsymbol{T})^{-1} \boldsymbol{e}.$$

□

The next is a recurring example used throughout this manuscript.

Example 2.1 *Fix $n \geq 1$ and $a > 0$. Let $X \sim \mathrm{PH}_n(\boldsymbol{\pi}, \boldsymbol{T})$ with*

$$\boldsymbol{\alpha} = (1, 0, \ldots, 0) \quad \textit{and} \quad \boldsymbol{T} = \begin{pmatrix} -n/a & n/a & 0 & \cdots & 0 \\ 0 & -n/a & n/a & \cdots & 0 \\ 0 & 0 & -n/a & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -n/a \end{pmatrix}.$$

Notice that the distribution of X corresponds to the convolution of n exponential distributions of intensity n/a, which implies that $\mathbb{E}(X) = a$. The random variable X is said to follow an **Erlang distribution of n stages and mean** *a. It can be verified that the density function f of X is given by*

$$f(x) = \frac{(n/a)^n}{(n-1)!} x^{n-1} e^{-nx/a}, \quad x \geq 0. \tag{2.2.2}$$

.

The following further showcases some helpful probabilistic arguments used in the theory of matrix–analytic methods.

Theorem 2.10 *Let $\tau_1 \sim \mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$ and $\tau_2 \sim \mathrm{PH}(\boldsymbol{\alpha}, \boldsymbol{S})$ be independent random variables with respective distribution functions denoted by F_{τ_1} and F_{τ_2}. Let $\boldsymbol{t} = -\boldsymbol{T}\boldsymbol{e}$ and $\boldsymbol{s} = -\boldsymbol{S}\boldsymbol{e}$. The following hold.*

1. *For any $p \in (0,1)$, $pF_{\tau_1} + (1-p)F_{\tau_2}$ corresponds to a phase–type distribution with parameters*
$$\left((p\boldsymbol{\pi}, (1-p)\boldsymbol{\alpha}), \begin{pmatrix} \boldsymbol{T} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S} \end{pmatrix}\right). \tag{2.2.3}$$
2. *$\tau_1 + \tau_2$ is phase–type–distributed with parameters*
$$\left((\boldsymbol{\pi}, \boldsymbol{0}), \begin{pmatrix} \boldsymbol{T} & \boldsymbol{t}\boldsymbol{\alpha} \\ \boldsymbol{0} & \boldsymbol{S} \end{pmatrix}\right). \tag{2.2.4}$$
3. *$\min(\tau_1, \tau_2)$ is phase–type–distributed with parameters*
$$(\boldsymbol{\pi} \otimes \boldsymbol{\alpha}, \boldsymbol{T} \oplus \boldsymbol{S}), \tag{2.2.5}$$
where $\otimes$ and $\oplus$ denote the Kronecker product and Kronecker sum of matrices, respectively (see Appendix A).
4. *$\max(\tau_1, \tau_2)$ is phase–type–distributed with parameters*
$$\left((\boldsymbol{\pi} \otimes \boldsymbol{\alpha}, \boldsymbol{0}, \boldsymbol{0}), \begin{pmatrix} \boldsymbol{T} \oplus \boldsymbol{S} & \boldsymbol{t} \otimes \boldsymbol{I} & \boldsymbol{I} \otimes \boldsymbol{s} \\ \boldsymbol{0} & \boldsymbol{S} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{T} \end{pmatrix}\right). \tag{2.2.6}$$

Proof. Let $\mathfrak{E}_T$ and $\mathfrak{E}_S$ be the phase–spaces of $\mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$ and $\mathrm{PH}(\boldsymbol{\alpha}, \boldsymbol{S})$, respectively. Let us study the Markov jump processes associated to (2.2.3), (2.2.4), (2.2.5) and (2.2.6).

1. The structure of (2.2.3) indicates that the process either starts in the first block of states $\mathfrak{E}_T$ with probability p, or in the second block $\mathfrak{E}_S$ with probability $1-p$. If it starts in the first block, it will do so according to the initial distribution $\boldsymbol{\pi}$ and will evolve within $\mathfrak{E}_T$ according to $\boldsymbol{T}$ before absorption. If it starts in the second block, it will do so according to the initial distribution $\boldsymbol{\alpha}$ and will evolve within $\mathfrak{E}_S$ according to $\boldsymbol{S}$ before absorption. Thus, the time spent in $\mathfrak{E}_T \cup \mathfrak{E}_S$ corresponds to the mixture of F_{τ_1} and F_{τ_2}.
2. The process with parameters (2.2.4) starts in the block of states $\mathfrak{E}_T$ according to $\boldsymbol{\pi}$, evolves there according to $\boldsymbol{T}$, and exits $\mathfrak{E}_T$ according to the absorption intensities vector $\boldsymbol{t}$, only to enter the block of states $\mathfrak{E}_S$ according to $\boldsymbol{\alpha}$, evolving there according to $\boldsymbol{S}$, and getting absorbed from there later. Thus, (2.2.4) describes the **concatenation** of the underlying Markov jump processes of τ_1 and τ_2; see Figure 2.3.

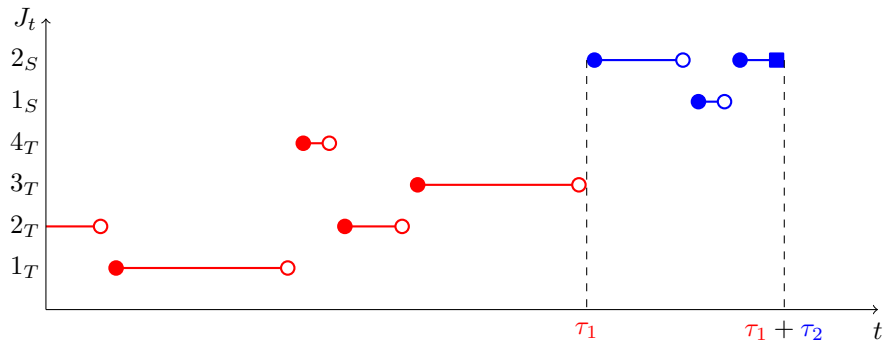


Figure 2.3: Concatenation of the underlying Markov jum processes of τ_1 and τ_2. The holding times shown in red correspond to the time spent in $\mathfrak{E}_T = \{1_T, 2_T, 3_T, 4_T\}$, and their sum correspond to τ_1. At time τ_1, the underlying process jumps to some state in $\mathfrak{E}_S = \{1_S, 2_S\}$. The holding times shown in blue correspond to the time spent in $\mathfrak{E}_S$, and their sum correspond to τ_2. At time $\tau_1 + \tau_2$, the underlying process exits the phase–type $\mathfrak{E}_T \cup \mathfrak{E}_S$ and gets absorbed.

3. Recall that by definition of the Kronecker product and Kronecker sum,

$$\boldsymbol{\pi} \otimes \boldsymbol{\alpha} = (\pi_1 \boldsymbol{\alpha}, \cdots, \pi_p \boldsymbol{\alpha}), \quad \text{and}$$

$$\begin{aligned} \boldsymbol{T} \oplus \boldsymbol{S} &= \boldsymbol{T} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{S} \\ &= \begin{pmatrix} t_{11}\boldsymbol{I} + \boldsymbol{S} & t_{12}\boldsymbol{I} & t_{13}\boldsymbol{I} & \cdots & t_{1p}\boldsymbol{I} \\ t_{21}\boldsymbol{I} & t_{22}\boldsymbol{I} + \boldsymbol{S} & t_{23}\boldsymbol{I} & \cdots & t_{2p}\boldsymbol{I} \\ t_{31}\boldsymbol{I} & t_{32}\boldsymbol{I} & t_{33}\boldsymbol{I} + \boldsymbol{S} & \cdots & t_{3p}\boldsymbol{I} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{p1}\boldsymbol{I} & t_{p2}\boldsymbol{I} & t_{p3}\boldsymbol{I} & \cdots & t_{pp}\boldsymbol{I} + \boldsymbol{S} \end{pmatrix}. \end{aligned} \tag{2.2.7}$$

Define $\mathfrak{E}_{T\times S} = \{(i,j) : i \in \mathfrak{E}_T, j \in \mathfrak{E}_S\}$, with its elements ordered lexicographically. With this in mind, the Markov jump process described by (2.2.5) is such that, its initial state, say $(i_0, j_0) \in \mathfrak{E}_{T\times S}$, is chosen with probability $\pi_{i_0}\alpha_{j_0}$, or in other words, the first entry is chosen according to $\boldsymbol{\pi}$ and the second one according to $\boldsymbol{\alpha}$, independently. Once in (i_0, j_0), the first jump can happen in one of the following three ways: it lands in some (k, j_0) with intensity $t_{i_0 k}$, it lands in some (i_0, l) with intensity $s_{j_0 l}$, or it gets absorbed with intensity $t_{i_0} + s_{j_0}$. Notice that jumps between states in $\mathfrak{E}_{T\times S}$ maintain one of their entries fixed, and the one that changes, will do it according to the intensities of $\boldsymbol{T}$ or $\boldsymbol{S}$, depending if it is the first entry or the second entry the one that changes. Absorption comes from the first entry (with intensity $\boldsymbol{t}$) or from the second entry (with intensity $\boldsymbol{s}$).

This is basically the same as letting two independent Markov jump process with parameters $(\boldsymbol{\pi}, \boldsymbol{T})$ and $(\boldsymbol{\alpha}, \boldsymbol{S})$ evolve in a parallel fashion; in the moment either of them gets absorbed, we stop both processes. Thus, the time until absorption of the process with parameters $(\boldsymbol{\pi} \otimes \boldsymbol{\alpha}, \boldsymbol{T} \oplus \boldsymbol{S})$ is the minimum between the times of absorption of parallel (and independent) Markov jump processes with parameters $(\boldsymbol{\pi}, \boldsymbol{T})$ and $(\boldsymbol{\alpha}, \boldsymbol{S})$.

4. The process associated to (2.2.6) is slightly more complicated than the one of (2.2.5). Its state–space is $\mathfrak{E}_{T\times S} \cup \mathfrak{E}_S \cup \mathfrak{E}_T$. Its initial distribution indicates that it starts in some state of $\mathfrak{E}_{T\times S}$, and it stays there until either the process associated to $(\boldsymbol{\pi}, \boldsymbol{T})$ or the process associated to $(\boldsymbol{\alpha}, \boldsymbol{S})$ gets absorbed. If the former happens, then it jumps to $\mathfrak{E}_S$ (according to $\boldsymbol{t} \otimes \boldsymbol{I}$) and stays there until the process $(\boldsymbol{\alpha}, \boldsymbol{S})$ gets absorbed too. If the latter happens, then it jumps to $\mathfrak{E}_T$ (according to $\boldsymbol{I} \otimes \boldsymbol{s}$) and stays there until the process $(\boldsymbol{\pi}, \boldsymbol{T})$ gets absorbed too. In any case, the process associated to (2.2.6) terminates whenever both of the processes associated to τ_1 and τ_2 finish.

□

We can further extend Theorem 2.10 to show that order statistics of observations sampled according to a phase–type distribution are also phase–type–distributed. More specifically, consider a sample of i.i.d.r.v.'s $\{\tau_1, \ldots, \tau_n\}$. We define the k-th order statistic of the sample of size n as the k-th smallest value amongst $\{\tau_1, \ldots, \tau_n\}$, which we denote by $\tau_{k:n}$. This means that $\tau_{1:n} \leq \tau_{2:n} \leq \cdots \leq \tau_{n:n}$. We claim that if $\{\tau_i\}_{i=1}^n$ are phase–type–distributed, then so are $\{\tau_{k:n}\}_{k=1}^n$. In order to prove this, we introduce the following notation.

Definition 2.11 Let $\boldsymbol{A}$ be a $m \times n$ matrix and let $\boldsymbol{B}$ be either a $p \times p$ or a $p \times 1$ matrix. For such matrices and $n \geq 1$, define

$$\boldsymbol{A}^{\otimes n} = \underbrace{\boldsymbol{A} \otimes \cdots \otimes \boldsymbol{A}}_{n \text{ terms}}.$$

$$\boldsymbol{B}^{\oplus n} = \sum_{j=0}^{n-1} \boldsymbol{I}^{\otimes j} \otimes \boldsymbol{B} \otimes \boldsymbol{I}^{\otimes n-1-j},$$

where $\boldsymbol{I}$ is of dimension $p \times p$.

Theorem 2.12 *Let* $\tau_1, \cdots, \tau_n$ *i.i.d.* $\sim \mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$.

1. *Then*

$$\min(\tau_1, \ldots, \tau_n) \sim \mathrm{PH}(\boldsymbol{\pi}^{\otimes n}, \boldsymbol{T}^{\oplus n}). \tag{2.2.8}$$

2. *If* $\boldsymbol{t} = -\boldsymbol{T}\boldsymbol{e}$, $\boldsymbol{\pi}_{k:n} = (\boldsymbol{\pi}^{\otimes n}, \mathbf{0}, \mathbf{0}, \ldots, \mathbf{0})$ *and*

$$\boldsymbol{T}_{k:n} = \begin{pmatrix} \boldsymbol{T}^{\oplus n} & \boldsymbol{t}^{\oplus(n)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{T}^{\oplus(n-1)} & \boldsymbol{t}^{\oplus(n-1)} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{T}^{\oplus(n-k+1)} \end{pmatrix},$$

then $\tau_{k:n} \sim \mathrm{PH}(\boldsymbol{\pi}_{k:n}, \boldsymbol{T}_{k:n})$.

Proof.

1. This is straightforward to verify using (2.2.5) and induction over n.
2. Consider the system of n independent Markov jump processes $\{J_i(t)\}_{t\geq 0}$ $(i = 1, \ldots, n)$ underlying $\tau_1, \ldots, \tau_n$; let $\mathfrak{E}_n$ denote the product state–space of a such system, ordered lexicographically. According to (2.2.8), the first time a Markov jump process gets absorbed is phase–type–distributed with sub–intensity matrix $\boldsymbol{T}^{\oplus n} = \boldsymbol{T} \oplus \cdots \oplus \boldsymbol{T}$ (n terms). Killing one process and letting the $n-1$ remaining continue to run is taken care of by the matrix $\boldsymbol{t}^{\oplus n}$, which makes a transition to $\mathfrak{E}_{n-1}$ where only the $n-1$ remaining processes are present. Repeating this argument k times will provide us with $\tau_{k:n}$ and the resulting structure of $\boldsymbol{T}_{k:n}$.

□

The following provides a recursive formula to compute moments of order statistics of a phase–type distribution.

Theorem 2.13 *Let* $\tau_1, \ldots, \tau_n$ *i.i.d.* $\sim \mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$ *and let* $\mu_{k:n} = \mathbb{E}(\tau_{k:n})$. *Then, for* $k = 1, \ldots, n-1$,

$$\mu_{k+1:n} = \mu_{k:n} + \boldsymbol{\pi}^{\otimes n} \left(\prod_{i=0}^{k-1} \left[-\boldsymbol{T}^{\oplus(n-i)} \right]^{-1} \boldsymbol{t}^{\oplus(n-i)} \right) \left(-\boldsymbol{T}^{\oplus(n-k)} \right)^{-1} \boldsymbol{e}, \quad (2.2.9)$$

where

$$\mu_{1:n} = \boldsymbol{\pi}^{\otimes n} \left(-\boldsymbol{T}^{\oplus n} \right)^{-1} \boldsymbol{e}.$$

Proof. The equality $\mu_{k:n} = \mathbb{E}(\tau_{k:n}) = \boldsymbol{\alpha}_{k:n}(-\boldsymbol{T}_{k:n})^{-1}\boldsymbol{e}$ and the block inversion formula

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \mathbf{0} & \boldsymbol{C} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{C}^{-1} \\ \mathbf{0} & \boldsymbol{C}^{-1} \end{pmatrix}$$

imply that

$$
\begin{aligned}
\mu_{k+1:n} &= \boldsymbol{\pi}_{k+1:n}(-\boldsymbol{T}_{k+1:n})^{-1}\boldsymbol{e} \\
&= (\boldsymbol{\pi}_{k:n}, \mathbf{0}) \begin{pmatrix} -\boldsymbol{T}_{k:n} & \begin{pmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ -\boldsymbol{t}^{\oplus(n-k+1)} \end{pmatrix} \\ \mathbf{0} & -\boldsymbol{T}^{\oplus(n-k)} \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{e} \\ \boldsymbol{e} \end{pmatrix} \\
&= (\boldsymbol{\pi}_{k:n}, \mathbf{0}) \begin{pmatrix} (-\boldsymbol{T}_{k:n})^{-1} & (-\boldsymbol{T}_{k:n})^{-1} \begin{pmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \boldsymbol{t}^{\oplus(n-k+1)} \end{pmatrix} (-\boldsymbol{T}^{\oplus(n-k)})^{-1} \\ \mathbf{0} & (-\boldsymbol{T}^{\oplus(n-k)})^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{e} \\ \boldsymbol{e} \end{pmatrix} \\
&= \boldsymbol{\pi}_{k:n}(-\boldsymbol{T}_{k:n})^{-1}\boldsymbol{e} + \boldsymbol{\pi}_{k:n}(-\boldsymbol{T}_{k:n})^{-1} \begin{pmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \boldsymbol{t}^{\oplus(n-k+1)} \end{pmatrix} (-\boldsymbol{T}^{\oplus(n-k)})^{-1}\boldsymbol{e}.
\end{aligned}
$$

Furthermore, since the inverse of the matrix

$$
\begin{pmatrix}
\boldsymbol{A}_1 & \boldsymbol{B}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\
\mathbf{0} & \boldsymbol{A}_2 & \boldsymbol{B}_2 & \mathbf{0} & \cdots & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \boldsymbol{A}_3 & \boldsymbol{B}_3 & \cdots & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{A}_4 & \cdots & \mathbf{0} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{A}_m
\end{pmatrix}
$$

is given by

$$
\begin{pmatrix}
\boldsymbol{A}_1^{-1} & \boldsymbol{A}_1^{-1}(-\boldsymbol{B}_1)\boldsymbol{A}_2^{-1} & \boldsymbol{A}_1^{-1}\boldsymbol{B}_1\boldsymbol{A}_2^{-1}\boldsymbol{B}_2\boldsymbol{A}_3^{-1} & \cdots & \left(\prod_{i=1}^{m-1}\boldsymbol{A}_i^{-1}(-\boldsymbol{B}_i)\right)\boldsymbol{A}_m^{-1} \\
\mathbf{0} & \boldsymbol{A}_2^{-1} & -\boldsymbol{A}_2^{-1}(-\boldsymbol{B}_2)\boldsymbol{A}_3^{-1} & \cdots & \left(\prod_{i=2}^{m-1}\boldsymbol{A}_i^{-1}(-\boldsymbol{B}_i)\right)\boldsymbol{A}_m^{-1} \\
\mathbf{0} & \mathbf{0} & \boldsymbol{A}_3^{-1} & \cdots & \left(\prod_{i=3}^{m-1}\boldsymbol{A}_i^{-1}(-\boldsymbol{B}_i)\right)\boldsymbol{A}_m^{-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{A}_m^{-1}
\end{pmatrix},
$$

then

$$
\begin{aligned}
&\boldsymbol{\pi}_{k:n}(-\boldsymbol{T}_{k:n})^{-1}\begin{pmatrix}\mathbf{0}\\ \vdots \\ \mathbf{0}\\ \boldsymbol{t}^{\oplus(n-k+1)}\end{pmatrix}\\
&\quad=(\boldsymbol{\pi}^{\otimes n},\mathbf{0},\ldots,\mathbf{0})\\
&\qquad\times\begin{pmatrix}-\boldsymbol{T}^{\oplus n} & -\boldsymbol{t}^{\oplus(n)} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & -\boldsymbol{T}^{\oplus(n-1)} & -\boldsymbol{t}^{\oplus(n-1)} & \cdots & \mathbf{0}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & -\boldsymbol{T}^{\oplus(n-k+1)}\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{0}\\ \vdots \\ \mathbf{0}\\ \boldsymbol{t}^{\oplus(n-k+1)}\end{pmatrix}\\
&\quad=\boldsymbol{\pi}^{\otimes n}\left\{\left(\prod_{i=1}^{k-1}\left[-\boldsymbol{T}^{\oplus(n-i+1)}\right]^{-1}\boldsymbol{t}^{\oplus(n-i+1)}\right)(-\boldsymbol{T}^{\oplus(n-k+1)})^{-1}\right\}\boldsymbol{t}^{\oplus(n-k+1)}\\
&\quad=\boldsymbol{\pi}^{\otimes n}\left(\prod_{i=0}^{k-1}\left[-\boldsymbol{T}^{\oplus(n-i)}\right]^{-1}\boldsymbol{t}^{\oplus(n-i)}\right),
\end{aligned}
$$

and the proof is finished. □

Finally, we prove the following result regarding the distribution of a geometric random sum of i.i.d. phase–type random variables.

Theorem 2.14 *Let* $\{\tau_i\}_{i=1}^{\infty}$ *be an i.i.d. sequence with common distribution* $\mathrm{PH}(\boldsymbol{\alpha},\boldsymbol{S})$ *and let* $Z\sim\mathrm{Geo}(1-\rho)$ *with support on* $\{0,1,2,\ldots\}$ *and* $\rho\in(0,1)$. *Then*

$$
\sum_{i=1}^{Z}\tau_i\sim\mathrm{PH}(\rho\boldsymbol{\alpha},\boldsymbol{S}+\rho\boldsymbol{s}\boldsymbol{\alpha}),\tag{2.2.10}
$$

which has a point mass at 0 *of size* $(1-\rho)$.

Proof. The underlying process $\{J_t^g\}_{t\geq 0}$ of the geometric random sum in (2.2.10) corresponds to a terminating sequential concatenation of the underlying processes $\{J_\cdot^i\}_{i\geq 1}$ of $\{\tau_i\}_{i\geq 1}$, where a given underlying process is added to the concatenation with probability ρ or termination happens with probability $1-\rho$. The memoryless property of the geometric distribution implies that

$\{J_t^g\}_{t\geq 0}$ is indeed an homogeneous Markov jump process. Since

$$\begin{aligned}
&\mathbb{P}(J_{\mathrm{d}t}^g = j \mid J_0^g = i) \\
&\quad = \mathbb{P}(J_{\mathrm{d}t}^g = j \mid J_0^1 = i) \\
&\quad = \mathbb{P}(J_{\mathrm{d}t}^1 = j \mid J_0^1 = i) + \mathbb{P}\left(\begin{array}{c} J_\cdot^1 \text{ terminates in } [0, \mathrm{d}t), J_0^2 = j, \\ \{\text{A second concatenation happens}\} \end{array} \,\middle|\, J_0^1 = i \right) \\
&\quad = (\delta_{ij} + s_{ij}\mathrm{d}t) + (s_i \mathrm{d}t)(\rho)(\alpha_j) \\
&\quad = \delta_{ij} + (\boldsymbol{S} + \rho \boldsymbol{s}\boldsymbol{\alpha})_{ij}\mathrm{d}t,
\end{aligned}$$

this means that the sub–intensity matrix of $\{J_t^g\}_{t\geq 0}$ is $\boldsymbol{S} + \rho\boldsymbol{s}\boldsymbol{\alpha}$. Since $\mathbb{P}(Z = 0) = 1 - \rho$, then $\{J_t^g\}_{t\geq 0}$ has a defective initial distribution $\rho\boldsymbol{\alpha}$ with a point mass at 0 of size $1 - \rho\boldsymbol{\alpha}\boldsymbol{e} = 1 - \rho$, so that (2.2.10) follows. □

2.2.2 Kulkarni's multivariate phase–type distributions

The first multivariate distribution with phase–type marginals (MPH) was proposed by Assaf et al. (1984). Their approach consisted in detecting the hitting times of different subsets of the state–space of a Markov jump process. This was later generalized in Kulkarni (1989), in which they considered vectors whose entries collect joint rewards from a single underlying Markov jump process; this class is commonly denoted by MPH*. A final generalization was proposed by Bladt and Nielsen (2010a), in which the authors considered vectors whose linear combination of entries are phase–type–distributed; this class is denoted by MVPH. Each one of these classes have their own advantages and disadvantages:

1. It is possible to explicitly compute multivariate densities from MPH, however, this class is somewhat limited from a stochastic modelling perspective.
2. The class MPH* is probabilistically appealing and flexible, however, there is no method to explicitly compute multivariate densities in MPH*; see Breuer (2016) for a semi–explicit method to compute bivariate MPH$*$ distributions.
3. MVPH is the most general class out of the three, however, aside from Bladt and Nielsen (2010a), little to no research has been done. This has to do with the fact that MVPH is defined in purely algebraic terms with no obvious probabilistic interpretation.

In this thesis we advocate for the use of MPH* distributions. Below we give a precise definition and a result which will be useful in forthcoming chapters.

Let $\{J_t\}_{t\geq 0}$ be a $(p+1)$–dimensional Markov jump process with intensity matrix

$$\begin{pmatrix} \boldsymbol{T} & \boldsymbol{t} \\ \boldsymbol{0} & 0 \end{pmatrix},$$

and initial vector $(\boldsymbol{\pi}, 0)$. Let $\boldsymbol{R} = \{r_{ij}\}$ be a $p \times n$–dimensional matrix with nonnegative entries. Consider the vector $(\tau_1, \ldots, \tau_n)$, where

$$\tau_i = \int_0^\infty r_{J_s i} \mathrm{d}s \quad \text{for all } i \in \{1, \ldots, n\}.$$

Then we say that $(\tau_1, \ldots, \tau_n)$ follows a **Kulkarni's multivariate phase–type distribution**, which we denote by $\mathrm{MPH}^*(\boldsymbol{\pi}, \boldsymbol{T}, \boldsymbol{R})$. See Figure 2.4 for the realization of a MPH^*–distributed random vector. In the following we prove

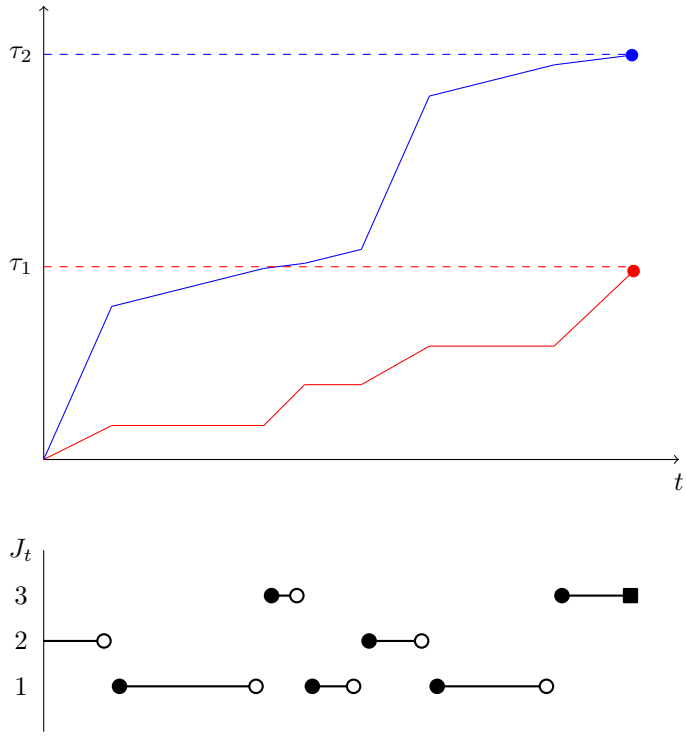


Figure 2.4: Realization of a MPH^*–distributed vector (τ_1, τ_2) with phase–space $\{1, 2, 3\}$ and rewards given by $r_{11} = 0$, $r_{21} = 0.5$, $r_{31} = 1$, and $r_{12} = 0.25$, $r_{22} = 2.25$, $r_{32} = 0.125$. The red path corresponds to $\int_0^t r_{J_s 1} \mathrm{d}s$ and the blue path corresponds to $\int_0^t r_{J_s 2} \mathrm{d}s$; the height of their values at the absorption time of $\{J_t\}_{t\geq 0}$ determines the value of τ_1 and τ_2, respectively.

that each τ_i indeed follows a phase–type distribution.

Theorem 2.15 *Let* $(\tau_1, \ldots, \tau_n) \sim \mathrm{MPH}^*(\boldsymbol{\pi}, \boldsymbol{T}, \boldsymbol{R})$. *Fix* $j \in \{1, \ldots, n\}$ *and w.l.o.g. suppose the row vector* $\boldsymbol{r}_j = \{r_{ij} : i \in \{1, \ldots, p\}\}$ *is on the form* $\boldsymbol{r}_j = (\boldsymbol{r}_j^+, \boldsymbol{0})$, *where* $\boldsymbol{r}_j^+ > 0$. *Let* $\boldsymbol{\Delta}_{r_j^+} = \mathrm{diag}\{\boldsymbol{r}_j^+\}$. *Partition the state space* $\{1, \ldots, p\}$ *into* $\mathfrak{E}_+ \cup \mathfrak{E}_0$, *where*

$$\mathfrak{E}_+ = \{i : r_{ij} > 0\} \quad \textit{and} \quad \mathfrak{E}_0 = \{i : r_{ij} = 0\},$$

and make a block partition of $\boldsymbol{\pi}$, $\boldsymbol{T}$ *and* $\boldsymbol{t} = -\boldsymbol{T}\boldsymbol{e}$ *accordingly:*

$$\boldsymbol{\pi} = (\boldsymbol{\pi}_+, \boldsymbol{\pi}_0), \qquad \boldsymbol{T} = \begin{pmatrix} \boldsymbol{T}_{++} & \boldsymbol{T}_{+0} \\ \boldsymbol{T}_{0+} & \boldsymbol{T}_{00} \end{pmatrix}, \qquad \boldsymbol{t} = \begin{pmatrix} \boldsymbol{t}_+ \\ \boldsymbol{t}_0 \end{pmatrix}.$$

1. *If* $\mathfrak{E}_0 = \emptyset$, *then*
$$\tau_j \sim \mathrm{PH}\left(\boldsymbol{\pi}, \boldsymbol{\Delta}_{r_j^+}^{-1}\boldsymbol{T}\right). \tag{2.2.11}$$
2. *If* $\mathfrak{E}_0 \neq \emptyset$, *then*
$$\tau_j \sim \mathrm{PH}\left(\boldsymbol{\pi}^c, \boldsymbol{T}^c\right), \tag{2.2.12}$$
where $\mathrm{PH}\left(\boldsymbol{\pi}^c, \boldsymbol{T}^c\right)$ *has an atom at* 0 *of size* $p_0 = \boldsymbol{\pi}_0(-\boldsymbol{T}_{00})^{-1}\boldsymbol{t}_0$, *and*
$$\boldsymbol{\pi}^c = \boldsymbol{\pi}_+ + \boldsymbol{\pi}_0(-\boldsymbol{T}_{00})^{-1}\boldsymbol{T}_{0+}, \quad \textit{and}$$
$$\boldsymbol{T}^c = \boldsymbol{\Delta}_{r_j^+}^{-1}(\boldsymbol{T}_{++} + \boldsymbol{T}_{+0}(-\boldsymbol{T}_{00})\boldsymbol{T}_{0+}).$$

Proof.

1. Suppose $\mathfrak{E}_0 = \emptyset$. For $i \in \{1, \ldots, p\}$, the contribution $\int_0^\infty r_{J_t j} \mathbb{1}_{J_t = i} \mathrm{d}t$ corresponds to the total occupation time of $\{J_t\}$ in state i scaled by a factor of r_{ij}. This is the same as "stretching" each holding time of $\{J_t\}$ in state i by a factor of r_{ij}: each one of those stretched holding times is $\mathrm{Exp}(-t_{ii}/r_{ij})$–distributed, with jump intensity to $k \neq i$ given by t_{ik}/r_{ij}. Repeating this argument for each $i \in \{1, \ldots, p\}$ we get that the sub–intensity matrix of the "stretched" underlying process of τ_j is given by $\boldsymbol{\Delta}_{r_j^+}^{-1}\boldsymbol{T}$ and the proof of (2.2.11) is finished.
2. If $\mathfrak{E}_0 \neq \emptyset$, then $\tau_j = 0$ if and only if $J_t \in \mathfrak{E}_0$ for all $t \geq 0$ before absorption. Thus,
$$\begin{aligned}
\mathbb{P}(\tau_j = 0) &= \sum_{i,k \in \mathfrak{E}_0} \int_0^\infty \mathbb{P}(J_0 = i)\mathbb{P}(J_r \in \mathfrak{E}_0 \text{ for all } r \leq s, J_s = k \mid J_0 = i) \\
&\qquad \times \mathbb{P}(J_{s+\mathrm{d}s} = p+1 \mid J_s = k) \\
&= \sum_{i,k \in \mathfrak{E}_0} \int_0^\infty (\boldsymbol{\pi}_0)_i \left(\exp(\boldsymbol{T}_{00}s)\right)_{ik} \left((\boldsymbol{t}_0)_k \mathrm{d}s\right) \\
&= \boldsymbol{\pi}_0(-\boldsymbol{T}_{00})^{-1}\boldsymbol{t}_0.
\end{aligned}$$

To study the case $\tau_j > 0$ we could use the same "stretching" arguments as in Item 1), however, we must first eliminate or censor the holding times of $\{J_t\}_{t\geq 0}$ while it is in $\mathfrak{E}_0$. See Figure 2.5 for a path with the $\mathfrak{E}_0$–states censored. We denote this $\mathfrak{E}_0$–censored process by $\{J_t^c\}_{t\geq 0}$, with $\{J_t^c\}_{t\geq 0}$

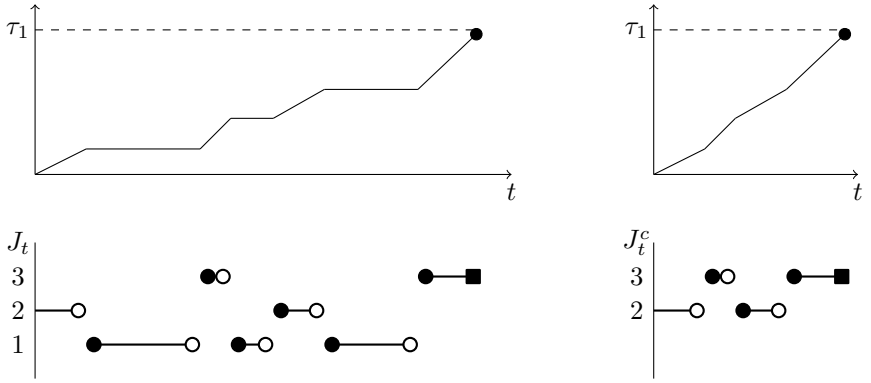


Figure 2.5: An example of a non–censored process underlying the marginal τ_1 (left) and its $\mathfrak{E}_0$–censored version (right), where $\mathfrak{E}_0 = \{1\}$ and $\mathfrak{E}_+ = \{2, 3\}$. The $\mathfrak{E}_0$–censored process $\{J_t^c\}_{t\geq 0}$ is obtained by eliminating the holding times of $\{J_t\}_{t\geq 0}$ at $\mathfrak{E}_0$. The height of the process $\int_0^t r_{J_t 1}\mathrm{d}s$ at the time of absorption of $\{J_t\}_{t\geq 0}$ coincides with the height of the process $\int_0^t r_{J_t^c 1}\mathrm{d}s$ at the time of absorption of $\{J_t^c\}_{t\geq 0}$, and both heights correspond to τ_1.

having state–space $\mathfrak{E}_+$. Let us compute the initial distribution $\boldsymbol{\pi}^c$ and sub–intensity matrix $\boldsymbol{T}^c$ of $\{J_t^c\}_{t\geq 0}$. First, for $i \in \mathfrak{E}_+$,

$$\begin{aligned}
\mathbb{P}(J_0^c = i) &= \mathbb{P}(J_0 = i) \\
&\quad + \sum_{k,\ell\in\mathfrak{E}_0} \int_0^\infty \begin{array}{r} \mathbb{P}(J_r \in \mathfrak{E}_0 \text{ for all } r \leq s, J_s = \ell \mid J_0 = k) \\ \times \mathbb{P}(J_0 = k)\mathbb{P}(J_{s+\mathrm{d}s} = i \mid J_s = \ell) \end{array} \\
&= (\boldsymbol{\pi}_+)_i + \sum_{k,\ell\in\mathfrak{E}_0} \int_0^\infty (\boldsymbol{\pi}_0)_k \left(\exp(\boldsymbol{T}_{00}s)\right)_{k\ell} \left((\boldsymbol{T}_{0+})_{\ell i}\mathrm{d}s\right) \\
&= (\boldsymbol{\pi}_+)_i + \left(\boldsymbol{\pi}_0(-\boldsymbol{T}_{00})^{-1}\boldsymbol{T}_{0+}\right)_i,
\end{aligned}$$

so that

$$\boldsymbol{\pi}^c = \boldsymbol{\pi}_+ + \boldsymbol{\pi}_0(-\boldsymbol{T}_{00})^{-1}\boldsymbol{T}_{0+}.$$

Also, for $i, k \in \mathfrak{E}_+$,

$$\begin{aligned}
&\mathbb{P}(J^c_{\mathrm{d}s} = k \mid J^c_0 = i) = \mathbb{P}(J^c_{\mathrm{d}s} = k, \mid J_0 = i) \\
&\quad = \mathbb{P}(J_{\mathrm{d}s} = k, \mid J_0 = i) \\
&\qquad + \sum_{\ell, h \in \mathfrak{E}_0} \mathbb{P}(J_{\mathrm{d}s} = \ell \mid J_0 = i) \int_{r=0}^{\infty} \begin{array}{c} \mathbb{P}(\{J_u\}_{u \le r} \subset \mathfrak{E}_0, J_u = h \mid J_{\mathrm{d}s} = \ell) \\ \times \mathbb{P}(J_{r+\mathrm{d}r} = k \mid J_r = h) \end{array} \\
&\quad = (\delta_{ik} + (\boldsymbol{T}_{++})_{ik}\mathrm{d}s) + \sum_{\ell h \in \mathfrak{E}_0} ((\boldsymbol{T}_{+0})_{i\ell}\mathrm{d}s) \int_{r=0}^{\infty} (\exp(\boldsymbol{T}_{00}r))_{\ell h} ((\boldsymbol{T}_{0+})_{hk}\mathrm{d}r) \\
&\quad = \delta_{ik} + \left(\boldsymbol{T}_{++} + \boldsymbol{T}_{+0}(-\boldsymbol{T}_{00})^{-1}\boldsymbol{T}_{0+}\right)_{ik} \mathrm{d}s,
\end{aligned}$$

so that

$$\boldsymbol{T}_c = \boldsymbol{T}_{++} + \boldsymbol{T}_{+0}(-\boldsymbol{T}_{00})^{-1}\boldsymbol{T}_{0+}.$$

Apply the arguments of Item 1) to the process $\{J^c_t\}$ with rewards $\boldsymbol{r}^+_j$ and (2.2.12) follows.

□

2.2.3 Matrix–exponential distributions

The class of matrix–exponential distributions was implicitly introduced in Cox (1955a) and Cox (1955b), where the author proposed to generalize the work of Jensen (1954) using hidden systems with "complex transition probabilities" instead. It was proved that this generalization coincides with the class of distributions on $[0, \infty)$ with rational Laplace transform, which by then had only been analysed by purely algebraic means. The systems proposed in Cox (1955a) and Cox (1955b) are defined through matrices and vectors with possibly complex entries, however, a direct probabilistic intepretation (in the classical probability sense) is not available as in the case of phase–type distributions. A survey of works related to matrix–exponential distributions, stemming from both the matrix–analytic point of view and from its algebraic point of view, can be found in Asmussen and O'Cinneide (2006). The following is a precise definition of a matrix–exponential distribution.

Definition 2.16 Let X be a nonnegative random variable with absolutely continuous distribution represented by a density f on the form

$$f(x) = \boldsymbol{\alpha} e^{\boldsymbol{S}x} \boldsymbol{s}, \qquad x \ge 0$$

for some p–dimensional complex–valued row vector $\boldsymbol{\alpha}$, $p \times p$ complex–valued matrix $\boldsymbol{S}$ and p–dimensional complex–valued column vector $\boldsymbol{S}$. Then we say that X follows a **matrix–exponential distribution of parameters** $(\boldsymbol{\alpha}, \boldsymbol{S}, \boldsymbol{s})$, denoted by either $\mathrm{ME}_p(\boldsymbol{\alpha}, \boldsymbol{S}, \boldsymbol{s})$ or $\mathrm{ME}(\boldsymbol{\alpha}, \boldsymbol{S}, \boldsymbol{s})$, depending if we want to be explicit about the dimension of its parameters or not. The parameters $(\boldsymbol{\alpha}, \boldsymbol{S}, \boldsymbol{s})$ will be regarded as the **representation** of the matrix–exponential distribution and p as the **dimension** of such representation.

Remark 3 *The representation of a given matrix–exponential distribution is non–unique. For instance, if $\boldsymbol{A}$ is a $p \times p$ non–singular matrix, then the parameters $(\boldsymbol{\alpha}\boldsymbol{A}, \boldsymbol{A}^{-1}\boldsymbol{S}\boldsymbol{A}, \boldsymbol{A}^{-1}\boldsymbol{s})$ and $(\boldsymbol{\alpha}, \boldsymbol{S}, \boldsymbol{s})$ yield the same matrix–exponential distribution. Indeed, for any $x \geq 0$,*

$$\begin{aligned}(\boldsymbol{\alpha}\boldsymbol{A})e^{\boldsymbol{A}^{-1}\boldsymbol{S}\boldsymbol{A}x}(\boldsymbol{A}^{-1}\boldsymbol{s}) &= (\boldsymbol{\alpha}\boldsymbol{A})\sum_{i=0}^{\infty}\frac{(\boldsymbol{A}^{-1}\boldsymbol{S}\boldsymbol{A}x)^i}{i!}(\boldsymbol{A}^{-1}\boldsymbol{s}) \\ &= (\boldsymbol{\alpha}\boldsymbol{A})\left(\boldsymbol{A}^{-1}\sum_{i=0}^{\infty}\frac{(\boldsymbol{S}x)^i}{i!}\boldsymbol{A}\right)(\boldsymbol{A}^{-1}\boldsymbol{s}) \\ &= \boldsymbol{\alpha}\sum_{i=0}^{\infty}\frac{(\boldsymbol{S}x)^i}{i!}\boldsymbol{s} = \boldsymbol{\alpha}e^{\boldsymbol{S}x}\boldsymbol{s}.\end{aligned}$$

Furthermore, a single matrix–exponential distribution can admit an infinite amount of representations of different dimensions. We say that a representation $(\boldsymbol{\alpha}, \boldsymbol{S}, \boldsymbol{s})$ of dimension p is **minimal** *if for any equivalent representation $(\boldsymbol{\alpha}_0, \boldsymbol{S}_0, \boldsymbol{s}_0)$ of dimension p_0, we have that $p \leq p_0$.*

Remark 4 *Theorem 2.7 implies that any phase–type distribution with parameters $(\boldsymbol{\pi}, \boldsymbol{T})$ corresponds to a matrix–exponential distribution of parameters $(\boldsymbol{\pi}, \boldsymbol{T}, \boldsymbol{t})$ with $\boldsymbol{t} = -\boldsymbol{T}\boldsymbol{e}$. However, there exist matrix–exponential distributions which are not phase–type; for a discussion on this, see O'Cinneide (1990), Maier (1991), Mocanu and Commault (1999), and Horváth and Telek (2015).*

The following are some well–known properties of matrix–exponential distributions (see Chapter 4 of Bladt and Nielsen (2017)).

- The Laplace transform of $X \sim \mathrm{ME}_p(\boldsymbol{\alpha}, \boldsymbol{S}, \boldsymbol{s})$ is on the form

$$L_X(\theta) = \mathbb{E}(e^{-\theta X}) = \frac{a_1\theta^{m-1} + \cdots + a_{m-1}\theta + a_m}{\theta^m + b_1\theta^{m-1} + \cdots + b_{m-1}\theta + b_m}, \quad \theta \geq 0 \quad (2.2.13)$$

  for some $m \leq p$ and real–valued constants $a_1, \ldots, a_m, b_1, \ldots, b_m$, where the r.h.s. of (2.2.13) is an irreducible fraction. Conversely, if a non–negative r.v. Y has a Laplace transform on the form (2.2.13) for some real–valued constants $a_1, \ldots, a_m, b_1, \ldots, b_m$, then Y is matrix–exponential–distributed.

- The poles of L_X, say $\{\kappa_i\}$ with respective multiplicities $\{m_i\}$, have strictly negative real part. Furthermore, there exists a unique pole with maximal real part, and such a pole is real.
- If f_X denotes the density function of $X \sim \mathrm{ME}_p(\boldsymbol{\alpha}, \boldsymbol{S}, \boldsymbol{s})$, then f_X (as a function of x) is an element of

$$\mathrm{span}\left\{\left\{\left\{x^{k-1}e^{\mathrm{Re}(\kappa_i)x}\cos(\mathrm{Im}(\kappa_i)x), x^{k-1}e^{\mathrm{Re}(\kappa_i)x}\sin(\mathrm{Im}(\kappa_i)x)\right\}_{k=1}^{m_i}\right\}_i\right\}, \tag{2.2.14}$$

  where $\{\kappa_i\}$ are the poles of L_X with respective multiplicities $\{m_i\}$. Conversely, if a density function f with support on $(0,\infty)$ is an element of (2.2.14), then f corresponds to the density function of a matrix–exponential distribution.
- The distribution of the random variable $X \sim \mathrm{ME}_p(\boldsymbol{\alpha}, \boldsymbol{S}, \boldsymbol{s})$ has an alternative representation $\mathrm{ME}_m(\boldsymbol{\alpha}^*, \boldsymbol{S}^*, \boldsymbol{s}^*)$ such that $\boldsymbol{\alpha}^*, \boldsymbol{S}^*$ and $\boldsymbol{s}^*$ have real entries, the eigenvalues of $\boldsymbol{S}^*$ correspond to the poles $\{\kappa_i\}$ of L_X with respective multiplicities $\{m_i\}$, and $\boldsymbol{s}^* = -\boldsymbol{S}^*\boldsymbol{e}$. This will be called a **minimal standard representation**. If we were to assume that $\mathrm{ME}_p(\boldsymbol{\alpha}, \boldsymbol{S}, \boldsymbol{s})$ is already a minimal and standard representation, we will denote it by $\mathrm{ME}_m(\boldsymbol{\alpha}, \boldsymbol{S})$. In most instances, from now on we will work with these type of representations.
- For a minimal standard representation $\mathrm{ME}_m(\boldsymbol{\alpha}, \boldsymbol{S})$,

$$\dim\left(\mathrm{span}\{\boldsymbol{\alpha}e^{\boldsymbol{S}x}, x \geq 0\}\right) = m. \tag{2.2.15}$$

While the class of matrix–exponential distribution shares some similarities with the class of phase–type distributions, their study is more algebraic–oriented rather than probabilistic. This is mainly due to the lack of a Markov jump process underlying the definition of a matrix–exponential distribution. However, not everything is lost: in O'Cinneide (1990) and O'Cinneide (1991a), a geometric understanding of the residual life of a matrix–exponential distribution is laid out. We will study this interpretation in depth in Subsection 2.3.2.

2.3 Point Processes

A large percentage of stochastic models have a stream of arrivals at random points in time, for instance, arrival of claims in a reserve risk process, or arrival of services in a queue. The mathematical framework to model these arrivals is through the theory of point processes. Below we give a precise definition of a point process and some additional elements associated to it.

Definition 2.17 Let $\{T_i\}_{i\geq 1}$ be real–valued random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$0 < T_1 < T_2 < T_3 < \cdots < \infty \quad \text{and} \quad \lim_{n\to\infty} T_n = \infty \quad \text{a.s..}$$

Define the random counting measure N over $(\mathbb{R}_+, \mathbb{B}(\mathbb{R}_+))$ by

$$N(A) = \sum_{i=1}^{\infty} \mathbb{1}\{T_i \in A\}, \quad A \in \mathbb{B}(\mathbb{R}_+).$$

We call N a **point process on** $\mathbb{R}_+$.

From the definition above, we call $\{T_i\}_{i\geq 1}$ the **arrival times** associated to N. For $t > 0$ denote $N_t = N((0,t])$. The stochastic process $\{N_t\}_{t\geq 0}$ completely determines $\{T_i\}_{i\geq 1}$ and the random measure N. We refer to $\{N_t\}_{t\geq 0}$ as the **arrival process**.

For all $s \geq 0$, denote by θ_s the **shift operator** defined by

$$\theta_s\{T_n\}_{n\geq 1} = \{T_{N_s+n} - s\}_{n\geq 1},$$

$$(\theta_s N)(A) = \sum_{i=1}^{\infty} \mathbb{1}\{T_{N_s+n} - s \in A\}, \quad A \in \mathbb{B}(\mathbb{R}_+), \qquad \text{and}$$

$$\theta_s N_t = (\theta_s N)((0,t]), \quad t \geq 0.$$

The space of realizations of the point process N is denoted by $\mathcal{N}$ and is given by

$$\mathcal{N} = \left\{ \nu : \begin{array}{c} \nu \text{ is a measure on } (\mathbb{R}_+, \mathbb{B}(\mathbb{R}_+)), \\ \nu(\{t\}) \in \{0,1\},\ \nu((0,t]) < \infty \text{ for all } t \geq 0 \end{array} \right\}.$$

A sequence $\{\nu_n\}_{n\geq 1}$ of elements of $\mathcal{N}$ is said to **converge vaguely** to $\nu \in \mathcal{N}$ if and only if

$$\int_0^{\infty} f(x)\nu_n(\mathrm{d}x) \to \int_0^{\infty} f(x)\nu(\mathrm{d}x) \quad \text{as} \quad n \to \infty$$

for all functions $f : \mathbb{R}_+ \to \mathbb{R}_+$ which are continuous with compact support. It can be shown that there exists a metric d on $\mathcal{N}$ that corresponds to the vague convergence on $\mathcal{N}$ (see Grandell (1977)). With such a metric d we define the **vague topology** on $\mathcal{N}$ and we let $\mathbb{B}(\mathcal{N})$ be the Borel σ–algebra determined by this topology. Thus, a point process N is a random element of $(\mathcal{N}, \mathbb{B}(\mathcal{N}))$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

In the following, we lay out the framework to work with two classes of point processes which are intimately connected with phase–type and matrix–exponential distributions: the Markovian arrival processes and the Rational arrival processes.

2.3.1 Markovian arrival processes

In Neuts (1979), the author introduces a point process with similar characteristic to those of the class of phase–type distributions. Such a point process, which was later coined as a Markovian arrival process, has an underlying Markov jump process with "two kinds of jumps". The jumps of the first kind in the underlying Markov jump process stay hidden, while the jumps of the second kind determine the arrival times of the Markovian arrival process. Just as in the case of phase–type distributions, this class of point processes showcases a great level of tractability and flexibility. In fact, any point process can be approximated arbitrarily "well" by a Markovian arrival process; see Asmussen and Koole (1993) for details. In the following we give a precise mathematical definition of the Markovian arrival process.

Let $\boldsymbol{C}$ and $\boldsymbol{D}$ be $m \times m$-matrices such that

1. $\boldsymbol{D}$ is a sub-intensity matrix,
2. $\boldsymbol{D}$ is a non-negative matrix, and
3. $\boldsymbol{C} + \boldsymbol{D}$ is an intensity matrix.

Let us consider an Markov jump process $\{Z_t\}_{t\geq 0}$ with state–space $\mathbb{N}\times\{1,\dots,m\}$ ordered lexicographically, with intensity matrix given by

$$\begin{pmatrix} \boldsymbol{C} & \boldsymbol{D} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \boldsymbol{C} & \boldsymbol{D} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \boldsymbol{C} & \boldsymbol{D} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{C} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

and initial distribution given by

$$(\boldsymbol{\alpha}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \dots).$$

For all $t \geq 0$, let

$$N_t = \pi_1(Z_t), \quad \text{and} \quad J_t := \pi_2(Z_t)$$

where π_i denotes the i-th projection mapping. We say that $\{N_t\}_{t\geq 0}$ is the **Markovian arrival process (or MAP) with parameters** $(\boldsymbol{\alpha}, \boldsymbol{C}, \boldsymbol{D})$ and that $\{J_t\}_{t\geq 0}$ is its **underlying or modulating Markov jump process**. We may denote the previous process in several ways, depending if we want to omit the initial distribution $\boldsymbol{\alpha}$ and/or if we want to omit the dimension of the parameters; thus, the Markovian arrival process with parameters $(\boldsymbol{\alpha}, \boldsymbol{C}, \boldsymbol{D})$

will be denoted by either $\mathrm{MAP}_m(\boldsymbol{\alpha}, \boldsymbol{C}, \boldsymbol{D})$, $\mathrm{MAP}_m(\boldsymbol{C}, \boldsymbol{D})$, $\mathrm{MAP}(\boldsymbol{\alpha}, \boldsymbol{C}, \boldsymbol{D})$ or $\mathrm{MAP}(\boldsymbol{C}, \boldsymbol{D})$, depending on the situation.

From the characterisation above we get that for $i \neq j$,

$$\begin{aligned}\mathbb{P}(J_{t+\mathrm{d}t} = j, N_{t+\mathrm{d}t} = N_t \mid J_t = i) &= c_{ij}\mathrm{d}t,\\ \mathbb{P}(J_{t+\mathrm{d}t} = j, N_{t+\mathrm{d}t} = N_t + 1 \mid J_t = i) &= d_{ij}\mathrm{d}t,\\ \mathbb{P}(J_{t+\mathrm{d}t} = j \mid J_t = i) &= (c_{ij} + d_{ij})\mathrm{d}t,\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}(J_{t+\mathrm{d}t} = i, N_{t+\mathrm{d}t} = N_t \mid J_t = i) &= 1 + c_{ii}\mathrm{d}t,\\ \mathbb{P}(J_{t+\mathrm{d}t} = i, N_{t+\mathrm{d}t} = N_t + 1 \mid J_t = i) &= d_{ii}\mathrm{d}t,\\ \mathbb{P}(J_{t+\mathrm{d}t} = i \mid J_t = i) &= 1 + (c_{ii} + d_{ii})\mathrm{d}t.\end{aligned}$$

In other words, the intensity matrix of $\{J_t\}_{t\geq 0}$ is $\boldsymbol{C} + \boldsymbol{D}$. Arrival epochs of $\{N_t\}_{t\geq 0}$ occur according to the intensities of the matrix $\boldsymbol{D}$; the jumps of $\{J_t\}_{t\geq 0}$ occuring due to the intensities of $\boldsymbol{C}$ do not originate any arrivals. While $J_t = i$, an arrival of $\{N_t\}_{t\geq 0}$ happens with intensity d_{ii} without any jumps occuring in $\{J_t\}_{t\geq 0}$. See Figure 2.6 for the realization of a Markovian arrival process $\{N_t\}_{t\geq 0}$.

Let $\{T_i\}_{i\geq 1}$ denote the arrival times associated to $\{N_t\}_{t\geq 0}$, that is,

$$T_i = \inf\{t > 0 : N_t = i\}, \quad i \geq 1.$$

Notice that $\mathbb{P}(J_t = i, N_t = 0) = (\boldsymbol{\alpha} e^{\boldsymbol{C}t})_i$. Indeed, if the process $\{Z_t\}_{t\geq 0}$ stays in $0 \times \{1, \ldots, m\}$ up to time t, it will evolve according to the sub–intensity matrix $\boldsymbol{C}$. Then,

$$\begin{aligned}&\mathbb{P}(T_1 \in [t, t+\mathrm{d}t], J_{t+\mathrm{d}t} = j)\\ &\quad= \sum_{i=1}^{m} \mathbb{P}(J_t = i, N_t = 0)\mathbb{P}(T_1 \in [t, t+\mathrm{d}t], J_{t+\mathrm{d}t} = j \mid J_t = i, N_t = 0)\\ &\quad= \sum_{i=1}^{m} (\boldsymbol{\alpha} e^{\boldsymbol{C}t})_i (d_{ij}\mathrm{d}t)\\ &\quad= (\boldsymbol{\alpha} e^{\boldsymbol{C}t}\boldsymbol{D})_j \mathrm{d}t,\end{aligned}$$

and so

$$\mathbb{P}(T_1 \in [t, t+\mathrm{d}t]) = \boldsymbol{\alpha} e^{\boldsymbol{C}t}\boldsymbol{D}\boldsymbol{e}\mathrm{d}t.$$

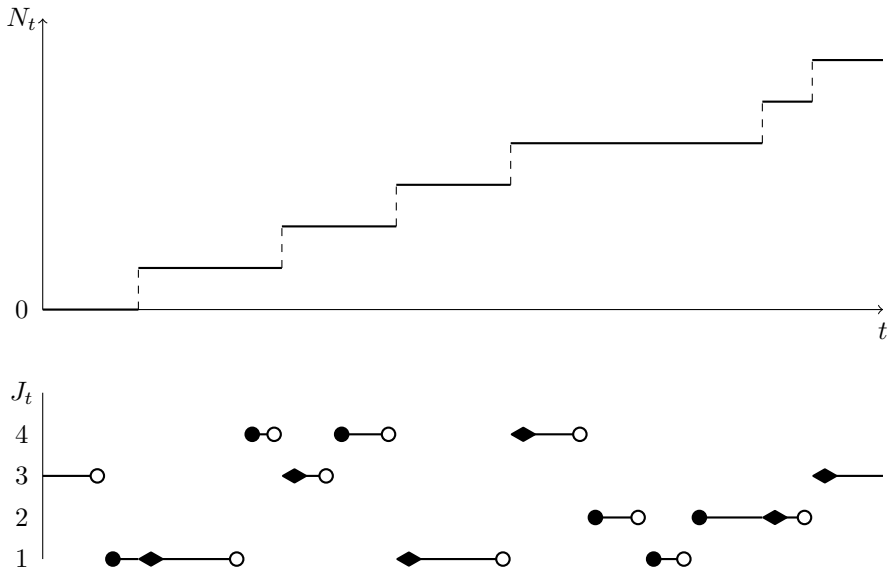


Figure 2.6: An example of a Markovian arrival process $\{N_t\}_{t\geq 0}$ with underlying Markov jump process $\{J_t\}_{t\geq 0}$. Arrival epochs which occur according to the intensities of the matrix $\boldsymbol{D}$ are marked with diamonds in $\{J_t\}_{t\geq 0}$. Jumps of $\{J_t\}_{t\geq 0}$ which occur according to the intenisty matrix $\boldsymbol{C}$ without originating arrivals in $\{N_t\}_{t\geq 0}$ are marked with circles.

Moreover,

$$\begin{aligned}\mathbb{P}(J_{T_1} = j) &= \int_0^\infty \mathbb{P}(T_1 \in [t, t+\mathrm{d}t], J_{t+\mathrm{d}t} = j) \\ &= \int_0^\infty (\boldsymbol{\alpha} e^{\boldsymbol{C}t}\boldsymbol{D})_j \mathrm{d}t \\ &= \left(\boldsymbol{\alpha}\left[\int_0^\infty e^{\boldsymbol{C}t}\mathrm{d}t\right]\boldsymbol{D}\right)_j \\ &= \left(\boldsymbol{\alpha}\left[-\boldsymbol{C}^{-1}\right]\boldsymbol{D}\right)_j.\end{aligned}$$

Either by iterating or by induction, we get the following.

Theorem 2.18 *Let $T_0 = 0$ and let $T_1, T_2, \ldots$ be the arrival times of the* $\mathrm{MAP}(\boldsymbol{\alpha}, \boldsymbol{C}, \boldsymbol{D})$ $\{N_t\}_{t\geq 0}$ *with underlying Markov jump process $\{J_t\}_{t\geq 0}$. Then, for all $n \geq 1$,*

$$\begin{aligned}&\mathbb{P}(T_1 \in [x_1, x_1 + \mathrm{d}x_1], T_2 - T_1 \in [x_2, x_2 + \mathrm{d}x_2] \ldots, T_n - T_{n-1} \in [x_n, x_n + \mathrm{d}x_n]) \\ &\quad = \boldsymbol{\alpha} e^{\boldsymbol{C}x_1}\boldsymbol{D}e^{\boldsymbol{C}x_2}\boldsymbol{D}\cdots e^{\boldsymbol{C}x_n}\boldsymbol{D}\boldsymbol{e}\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n, \quad \textit{and}\end{aligned}$$

$$T_n - T_{n-1} \sim \mathrm{PH}(\boldsymbol{\alpha}(-\boldsymbol{C}^{-1}\boldsymbol{D})^{n-1}, \boldsymbol{C}).$$

Below we give a couple of examples of Markovian arrival processes.

Example 2.2 (Poisson process) *Let $\lambda > 0$ and consider the Markovian arrival process $\{N_t\}_{t\geq 0}$ with parameters $\boldsymbol{C} = (-\lambda)$ and $\boldsymbol{D} = (\lambda)$. Then, for $x_1, \dots, x_n \geq 0$,*

$$\begin{aligned}&\mathbb{P}(T_1 \in [x_1, x_1 + \mathrm{d}x_1], T_2 - T_1 \in [x_2, x_2 + \mathrm{d}x_2] \dots, T_n - T_{n-1} \in [x_n, x_n + \mathrm{d}x_n]) \\ &\quad = (\lambda e^{-\lambda x_1})(\lambda e^{-\lambda x_2}) \cdots (\lambda e^{-\lambda x_n}) \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_n,\end{aligned}$$

so that $\{N_t\}_{t\geq 0}$ corresponds to a Poisson process of intensity λ.

Example 2.3 (PH–renewal process) *Let $(\boldsymbol{\pi}, \boldsymbol{T})$ correspond to the parameters of a phase–type distribution. Let $\{N_t\}_{t\geq 0}$ be a Markovian arrival process with parameters $(\boldsymbol{\pi}, \boldsymbol{T}, \boldsymbol{t}\boldsymbol{\pi})$. Then, for $x_1, \dots, x_n \geq 0$,*

$$\begin{aligned}&\mathbb{P}(T_1 \in [x_1, x_1 + \mathrm{d}x_1], T_2 - T_1 \in [x_2, x_2 + \mathrm{d}x_2] \dots, T_n - T_{n-1} \in [x_n, x_n + \mathrm{d}x_n]) \\ &\quad = \boldsymbol{\pi} e^{\boldsymbol{T}x_1}(\boldsymbol{t}\boldsymbol{\pi}) e^{\boldsymbol{T}x_2}(\boldsymbol{t}\boldsymbol{\pi}) \cdots e^{\boldsymbol{T}x_n}(\boldsymbol{t}\boldsymbol{\pi}) \boldsymbol{e} \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_n \\ &\quad = (\boldsymbol{\pi} e^{\boldsymbol{T}x_1} \boldsymbol{t})(\boldsymbol{\pi}) e^{\boldsymbol{T}x_2} \boldsymbol{t}) \cdots (\boldsymbol{\pi} e^{\boldsymbol{T}x_n} \boldsymbol{t}) \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_n,\end{aligned}$$

so that $\{N_t\}_{t\geq 0}$ corresponds to a renewal process with $\mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$–distributed interarrival times.

2.3.2 Rational arrival processes

In Asmussen and Bladt (1999), the authors introduced a point process based on the findings of O'Cinneide (1990) regarding the residual life of a matrix–exponential distribution. Such a point process, coined Rational arrival process, corresponds to the class of point processes with finite–dimensional–generated residual life. They proved that such a point process has an underlying PDMP called orbit process which provides an analogue to the Markov jump process underlying the Markovian arrival process. Furthermore, they showed that a Rational arrival process is characterised by a pair of matrices, and that its distribution is somewhat algebraically equivalent to the distribution of a Markovian arrival process. In the following we give a precise definition of the Rational arrival process and characterise the PDMP underlying it; throughout this subsection we follow closely the proofs and arguments of Asmussen and Bladt (1999) and Section 10.5 in Bladt and Nielsen (2017).

Let N be a point process defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Denote by $\mathscr{M}(\mathscr{N})$ the class of all finite signed measures on $(\mathscr{N}, \mathbb{B}(\mathscr{N}))$, the space of realizations of N. Let $\{N_t\}_{t\geq 0}$ be the arrival process associated to N and for all $t \geq 0$, define $\mathcal{F}_t = \sigma(\{N_s\}_{s\leq t})$. Write $\mu(t, \cdot)$ for a version of $\mathbb{P}(\theta_t N \in \cdot | \mathcal{F}_t)$ and define the measure $\mu^*(t, \omega)(\cdot) := \mu(t, \cdot)(\omega)$ for all $\omega \in \Omega$.

Definition 2.19 We call a point process N a **Rational arrival process** (RAP) if there exists a finite–dimensional subspace $V \subset \mathscr{M}(\mathscr{N})$, such that for any $t \geq 0$, $\mathbb{P}(\theta_t N \in \cdot | \mathcal{F}_t)$ has a version $\mu(t, \cdot)$ with $\mu^*(t, \omega) \in V$ for all $\omega \in \Omega$, and if $N((0, \infty)) = \infty$.

Remark 5 *Although the definition above is precise, most of its components are abstract and difficult to manipulate. The overall goal of this subsection is to give an alternative characterisation of the RAP in terms of certain PDMP, which is states in Proposition 2.27. Such characterisation will make the study of RAPs eaiser, and it will show that RAPs are an extension of the MAPs in a similar sense that matrix–exponential distributions are an extension of phase–type distributions.*

The Hahn–Jordan decomposition theorem (see Theorem 6.14 in Rudin (2006)) implies that if V is a finite–dimensional space of signed measures, we can choose a collection $v_1, \ldots, v_p$ of linearly independent probability measures that span V. In the following we show that for a RAP $\{N_t\}_{t \geq 0}$, there exists a compact set $V_c \subset V$ such that $\mu^*(t, \omega) \in V_c$ for all $t \geq 0$, $\omega \in \Omega$.

Lemma 2.20 *Let $v_1, \ldots, v_p$ be linearly independent probability measures that span V. Then*

$$\sup \left\{ \|\boldsymbol{a}\|_{\max} : \boldsymbol{a} = (a_1, \ldots, a_p) \in \mathbb{R}^p, \sum_{i=1}^{p} a_i v_i \text{ is a probability measure} \right\} < \infty.$$

Proof. Let $\boldsymbol{v} = (v_1, \cdots, v_p)'$. Assume there exists a sequence $\{\boldsymbol{a}^{(n)}\}_{n \geq 0}$ of elements in $\mathbb{R}^p$ such that $\mu^{(n)} := \boldsymbol{a}^{(n)} \boldsymbol{v}$ is a probability measure for each $n \geq 0$, and $\|\boldsymbol{a}^{(n)}\|_{\max} \to \infty$ as $n \to \infty$. The set $U = \{\boldsymbol{x} \in \mathbb{R}^p : \|\boldsymbol{x}\|_{\max} = 1\}$ is compact, so we can chose a subsequence $\{n_k\}$ such that $\boldsymbol{a}^{(n_k)} / \|\boldsymbol{a}^{(n_k)}\|_{\max} \to \boldsymbol{a}$ as $k \to \infty$ for some $\boldsymbol{a} \in U$. Then,

$$\boldsymbol{a}\boldsymbol{v} = \lim_{k \to \infty} \left(\frac{\boldsymbol{a}^{(n_k)}}{\|\boldsymbol{a}^{(n_k)}\|_{\max}} \right) \boldsymbol{v} = \lim_{k \to \infty} \frac{\boldsymbol{a}^{(n_k)} \boldsymbol{v}}{\|\boldsymbol{a}^{(n_k)}\|_{\max}} = 0,$$

contradicting the linear independence of the elements of $\boldsymbol{v}$. □

From now on, assume that the subspace V in Definition 2.19 is taken to be **minimal**, that is, if there exist a subspace $V' \subset \mathscr{M}(\mathscr{N})$ such that $\mu^*(t, \omega) \in V'$ for all $t \geq 0$, $\omega \in \Omega$, then $V \subseteq V'$. Fix a base of $\boldsymbol{v}$ of V as in Lemma 2.20. Define the $\mathcal{F}_t$–adapted $\mathbb{R}^p$–valued process $\{\boldsymbol{A}(t)\}_{t \geq 0}$ by taking

$$\mathbb{P}(\theta_t N \in \cdot | \mathcal{F}_t) = \sum_{i=1}^{p} A_i(t) v_i = \boldsymbol{A}(t) \boldsymbol{v};$$

the process $\{\boldsymbol{A}(t)\}_{t\geq 0}$ is well–defined by the linear independence of the elements of $\boldsymbol{v}$. The state–space of $\{\boldsymbol{A}(t)\}_{t\geq 0}$ is

$$\mathfrak{B} = \left\{ a \in \mathbb{R}^p : \boldsymbol{a}\boldsymbol{e} = 1, \sum_{i=1}^{p} a_i v_i(F) \geq 0 \text{ for all } F \in \mathbb{B}(\mathscr{N}) \right\},$$

which by Lemma 2.20 is compact and convex.

Proposition 2.21 *1. There exists a $p \times p$ matrix $\boldsymbol{Q}$ such that*

$$\boldsymbol{v} \circ \theta_s := \begin{pmatrix} v_1(\theta_s N \in \cdot) \\ \vdots \\ v_p(\theta_s N \in \cdot) \end{pmatrix} = e^{\boldsymbol{Q}s}\boldsymbol{v}.$$

2. $\boldsymbol{Q}\boldsymbol{e} = \boldsymbol{0}$ and $\mathrm{dev}(\boldsymbol{Q}) = 0$, where $\mathrm{dev}(\boldsymbol{Q})$ corresponds to the eigenvalue of $\boldsymbol{Q}$ with maximal real part.

3. $\mathbb{E}(\boldsymbol{A}(t+s)|\mathcal{F}_t) = \boldsymbol{A}(t)e^{\boldsymbol{Q}s}$. Equivalently, $\{\boldsymbol{A}(t)e^{-\boldsymbol{Q}t}\}_{t\geq 0}$ is a vector-valued martingale.

Proof.

1. We have that

$$\mathbb{P}(\theta_{t+s}N \in \cdot|\mathcal{F}_t) = \mathbb{P}(\theta_t\theta_s N \in \cdot|\mathcal{F}_t) = \sum_{i=1}^{p} A_i(t) v_i \circ \theta_s.$$

On the other hand,

$$\begin{aligned}\mathbb{P}(\theta_{t+s}N \in \cdot|\mathcal{F}_t) &= \mathbb{E}(\mathbb{P}(\theta_{t+s}N \in \cdot|\mathcal{F}_{t+s})|\mathcal{F}_t) \\ &= \sum_{i=1}^{p} \mathbb{E}(A_i(t+s)|\mathcal{F}_t) v_i.\end{aligned}$$

Fix $t \geq 0$. Then, for almost all $\omega \in \Omega$,

$$\boldsymbol{A}(t,\omega)(\boldsymbol{v} \circ \theta_s) = \boldsymbol{B}(t,s,\omega)\boldsymbol{v}, \quad \text{for all } s \geq 0, \tag{2.3.1}$$

where $\boldsymbol{B}(t,s,\cdot)$ is a version of $\mathbb{E}(\boldsymbol{A}(t+s)|\mathcal{F}_t)$. Now, let $t_1, \ldots, t_p > 0$ and choose $\omega_1, \ldots, \omega_p \in \Omega$ such that (2.3.1) holds for each pair (t_i, ω_i), and such that $\boldsymbol{A}(t_1,\omega_1), \ldots, \boldsymbol{A}(t_p,\omega_p)$ are linearly independent: if the latter were not possible, minimality of V would be contradicted. Thus

$$\boldsymbol{v} \circ \theta_s = \begin{pmatrix} \boldsymbol{A}(t_1,\omega_1) \\ \vdots \\ \boldsymbol{A}(t_p,\omega_p) \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{B}(t_1,s,\omega_1) \\ \vdots \\ \boldsymbol{B}(t_p,s,\omega_p) \end{pmatrix} \boldsymbol{v},$$

so that $\boldsymbol{v} \circ \theta_s = \tilde{Q}(s)\boldsymbol{v}$ for some matrix $\tilde{Q}(s)$ which is unique by linear independence of $\boldsymbol{v}$. Furthermore, by the semigroup property of θ_s, we have that $\tilde{Q}(s_1 + s_2) = \tilde{Q}(s_1)\tilde{Q}(s_2)$ for all $s_1, s_2 \geq 0$, and since $\boldsymbol{v} \circ \theta_s$ is right–continuous in s, we conclude that $\tilde{Q}(s) = e^{\boldsymbol{Q}s}$ for some $\boldsymbol{Q}$.

2. Notice that
$$e^{\boldsymbol{Q}s}\boldsymbol{e} = e^{\boldsymbol{Q}s}\boldsymbol{v}(\mathscr{N}) = \boldsymbol{v} \circ \theta_s(\mathscr{N}) = \boldsymbol{e},$$
so that
$$\frac{\mathrm{d}}{\mathrm{d}s} e^{\boldsymbol{Q}s}\boldsymbol{e}\Big|_{s=0} = \boldsymbol{Q}e^{\boldsymbol{Q}s}\boldsymbol{e}\big|_{s=0} = \boldsymbol{Q}\boldsymbol{e} = \boldsymbol{0}.$$
and $0 \in \mathrm{sp}(\boldsymbol{Q})$.
Now, suppose that there exists $a + \mathrm{i}b \in \mathrm{sp}(\boldsymbol{Q})$ and corresponding eigenvector $\boldsymbol{u} + \mathrm{i}\boldsymbol{w}$ such that $a > 0$. W.l.o.g, suppose that $\boldsymbol{u} \neq \boldsymbol{0}$. The linear independence of $\boldsymbol{v}$ implies that there exists $B \in \mathbb{B}(\mathscr{N})$ such that $\boldsymbol{u}\boldsymbol{v}(B) \neq 0$. Moreover, since $v_1, \ldots, v_p$ are probability measures, then
$$|\boldsymbol{u}(\boldsymbol{v} \circ \theta_s(B))| \leq \sum_i |u_i|, \quad \text{for all} \quad s \geq 0. \tag{2.3.2}$$
On the other hand, following the same steps that lead to (2.2.1) we have that
$$\begin{aligned}\boldsymbol{u}(\boldsymbol{v} \circ \theta_s(B)) &= \boldsymbol{u}e^{\boldsymbol{Q}s}\boldsymbol{v}(B) \\ &= e^{as}[\cos(bs)\boldsymbol{u}\boldsymbol{v}(B) - \sin(bs)\boldsymbol{w}\boldsymbol{v}(B)],\end{aligned}$$
which clearly is not bounded on s and contradicts (2.3.2). Thus, $\mathrm{dev}(\boldsymbol{Q}) = 0$.

3. We have that
$$\sum_{i=1}^{p} \mathbb{E}(A_i(t+s)|\mathcal{F}_t)v_i = \sum_{i=1}^{p} A_i(t)v_i \circ \theta_s = \sum_{i,j=1}^{p} A_i(t)(e^{\boldsymbol{Q}s})_{ji}v_j,$$
so that $\mathbb{E}(\boldsymbol{A}(t+s)|\mathcal{F}_t)\boldsymbol{v} = \boldsymbol{A}(t)e^{\boldsymbol{Q}s}\boldsymbol{v}$, which by linear independence of $\boldsymbol{v}$ implies $\mathbb{E}(\boldsymbol{A}(t+s)|\mathcal{F}_t) = \boldsymbol{A}(t)e^{\boldsymbol{Q}s}$.

□

Corollary 2.22 *The process $\{\boldsymbol{A}(t)\}_{t\geq 0}$ has a version with cádlág paths.*

Proof. This follows by a standard regularisation of the martingale $\{\boldsymbol{A}(t)e^{-\boldsymbol{Q}t}\}_{t\geq 0}$; see Rogers and Williams (1993) for details. □

For all $t \geq 0$, define the operator $R_t : \mathscr{M}(\mathscr{N}) \to \mathscr{M}(\mathscr{N})$ by taking

$$R_t\mu(A) = \mu\{N \in \mathscr{N} : \theta_t N \in A, T_1 > t\}, \quad \mu \in \mathscr{M}(\mathscr{N}), A \in \mathbb{B}(\mathbb{R}_+),$$

so that

$$R_t\mathbb{P}(N \in \cdot) = \mathbb{P}(\theta_t N \in \cdot, T_1 > t).$$

We will call R_t the **residual life operator**. Below we prove some of its properties.

Proposition 2.23 *1.* $\{R_t\}_{t\geq 0}$ *is a semigroup.*
2. $\mu \in V \Rightarrow R_t\mu \in V$.

Proof.

1. Let $G = \{N \in \mathscr{N} : \theta_t N \in A, T_1(N) > t\}$. Then

$$\begin{aligned} R_s R_t\mu(A) &= R_s\mu(G) \\ &= \mu\{N \in \mathscr{N} : \theta_s N \in G, T_1(N) > s\} \\ &= \mu\{N \in \mathscr{N} : \theta_t\theta_s N \in A, T_1(\theta_s N) > t, T_1(N) > s\} \\ &= \mu\{N \in \mathscr{N} : \theta_{s+t} N \in A, T_1(N) > t + s\} \\ &= R_{s+t}\mu(A). \end{aligned}$$

2. If $\mu \in \mathscr{M}(\mathscr{N})$, then

$$\begin{aligned} R_t\mu(\cdot) &= R_t\mathbb{P}(N \in \cdot) \\ &= \mathbb{P}(\theta_t N \in \cdots, T_1 > t) \\ &= \mathbb{E}(\mathbb{P}(\theta_t N \in \cdot, T_1 > t|\mathcal{F}_t)) \\ &= \mathbb{E}(\mathbb{1}_{T_1>t}\mathbb{P}(\theta_t N \in \cdot|\mathcal{F}_t)) \\ &= \mathbb{E}\left(\sum_{i=1}^{p} \mathbb{1}_{T_1>t} A_i(t) v_i(\cdot)\right) \\ &= \sum_{i=1}^{p} \mathbb{E}(\mathbb{1}_{T_1>t} A_i(t)) v_i(\cdot) \in V. \end{aligned}$$

□

Theorem 2.24 *There exists a matrix* $\boldsymbol{C}$ *such that*

$$R_t(\boldsymbol{a}\boldsymbol{v}) = \boldsymbol{a}e^{\boldsymbol{C}t}\boldsymbol{v} \quad \textit{for all} \quad \boldsymbol{a} \in \mathfrak{B}. \tag{2.3.3}$$

Furthermore, $\mathrm{dev}(\boldsymbol{C}) < 0$.

PROOF. Equation (2.3.3) follows by the semigroup property of $\{R_t\}_{t\geq 0}$; see Section 4.5 in Bladt and Nielsen (2017) for technical details. Since $\mathbb{P}(T_1 > t) \to 0$ as $t \to \infty$, it follows that $R_t(\boldsymbol{a}\boldsymbol{v}) = \boldsymbol{a}e^{\boldsymbol{C}t}\boldsymbol{v} \to 0$. The linear independence of $v_1, \ldots, v_p$ implies that $\boldsymbol{a}e^{\boldsymbol{C}t} \to \boldsymbol{0}$, and so, minimality of $\mathfrak{B}$ and Theorem 2.8 imply that $\mathrm{Re}(\mathrm{dev}(\boldsymbol{C})) < 0$. That $\mathrm{dev}(\boldsymbol{C}) \in \mathbb{R}$ follows by noticing that if $\mathrm{dev}(\boldsymbol{C})$ were in $\mathbb{C} \setminus \mathbb{R}$, then the mapping $x \to \boldsymbol{a}e^{\boldsymbol{C}t}\boldsymbol{e}$ would downcross 0 an infinite number of times, which is a contradiction; see Section 5.6 of Fackrell (2003) for details. □

Lemma 2.25 *For all $t, h \geq 0$,*

$$\mathbb{P}(\theta_{t+h}N \in \cdot, N((t, t+h]) = 0|\mathcal{F}_t) = \boldsymbol{A}(t)e^{\boldsymbol{C}h}\boldsymbol{v}.$$

PROOF. This follows from

$$\mathbb{P}(\theta_{t+h}N \in \cdot, N((t, t+h]) = 0|\mathcal{F}_t) = R_h\mathbb{P}(\theta_t N \in \cdot|\mathcal{F}_t),$$

and Theorem 2.24. □

Proposition 2.26 *For all $t, h \geq 0$,*

$$\mathbb{P}(\theta_{t+h}N \in \cdot, N((t, t+h]) > 0|\mathcal{F}_t) = \boldsymbol{A}(t)\boldsymbol{D}\boldsymbol{v}h + o(h),$$

where $\boldsymbol{D} = \boldsymbol{Q} - \boldsymbol{C}$. In particular,

$$\mathbb{P}(N((t, t+h]) > 0|\mathcal{F}_t) = \boldsymbol{A}(t)\boldsymbol{D}\boldsymbol{e}h + o(h).$$

PROOF. Notice that

$$\begin{aligned}
&\mathbb{P}(\theta_{t+h}N \in \cdot, N((t, t+h]) > 0|\mathcal{F}_t) \\
&\quad = \mathbb{P}(\theta_{t+h}N \in \cdot|\mathcal{F}_t) - \mathbb{P}(\theta_{t+h}N \in \cdot, N((t, t+h]) = 0|\mathcal{F}_t) \\
&\quad = \boldsymbol{A}(t)e^{\boldsymbol{Q}h}\boldsymbol{v} - \boldsymbol{A}(t)e^{\boldsymbol{C}h}\boldsymbol{v} \\
&\quad = \boldsymbol{A}(t)(\boldsymbol{I} + h\boldsymbol{Q} + o(h) - \boldsymbol{I} - h\boldsymbol{C} + o(h))\boldsymbol{v},
\end{aligned}$$

and the proof is finished. □

Proposition 2.27 *The process $\{\boldsymbol{A}(t)\}_{t\geq 0}$ satisfies the equation*

$$\boldsymbol{A}(t) = \boldsymbol{A}(0) + \int_0^t \{\boldsymbol{A}(s)\boldsymbol{C} - \boldsymbol{A}(s)\boldsymbol{C}\boldsymbol{e}\cdot\boldsymbol{A}(s)\}\mathrm{d}s + \sum_{i:T_i\leq t}\left\{\frac{\boldsymbol{A}(T_i^-)\boldsymbol{D}}{\boldsymbol{A}(T_i^-)\boldsymbol{D}\boldsymbol{e}} - \boldsymbol{A}(T_i^-)\right\}, \tag{2.3.4}$$

where $\{T_i\}_{i\geq 1}$ corresponds to the sequence of arrival epochs of N.

Proof. First, let us show that

$$\mathbb{1}_{N((t,t+h])=0}\boldsymbol{A}(t+h) = \mathbb{1}_{N((t,t+h])=0}\frac{\boldsymbol{A}(t)e^{\boldsymbol{C}h}}{\boldsymbol{A}(t)e^{\boldsymbol{C}h}\boldsymbol{e}} \quad \text{a.s..} \tag{2.3.5}$$

By the linear independence of $\boldsymbol{v}$, this is equivalent to prove that the random variables

$$Z_1^B := \mathbb{P}(\theta_{t+h}N \in B, N((t,t+h]) = 0|\mathcal{F}_{t+h}), \quad \text{and}$$
$$Z_2^B := \mathbb{1}_{N((t,t+h])=0}\frac{\boldsymbol{A}(t)e^{\boldsymbol{C}h}\boldsymbol{v}(B)}{\boldsymbol{A}(t)e^{\boldsymbol{C}h}\boldsymbol{e}}$$

are equal (a.s.) for all $B \in \mathbb{B}(\mathcal{N})$. Keep $B \in \mathbb{B}(\mathcal{N})$ fixed. Then, by definition of conditional expectation and with the aid of the classic Dynkin's $\pi-\lambda$ theorem, it suffices to prove that

$$\mathbb{E}(Z_1^B\mathbb{1}_E\mathbb{1}_F) = \mathbb{E}(Z_2^B\mathbb{1}_E\mathbb{1}_F), \tag{2.3.6}$$

for all $E \in \mathcal{F}_t$ and $F \in \sigma(\{N_s\}_{s=t}^{t+h})$. However, for any given $F \in \sigma(\{N_s\}_{s=t}^{t+h})$ it either happens that $\mathbb{1}_{N((t,t+h])=0}\mathbb{1}_F = 0$ or $\mathbb{1}_{N((t,t+h])=0}\mathbb{1}_F = \mathbb{1}_{N((t,t+h])=0}$ a.s.. Thus, it suffices to prove that

$$\mathbb{E}(Z_1^B\mathbb{1}_E) = \mathbb{E}(Z_2^B\mathbb{1}_E),$$

for all $E \in \mathcal{F}_t$. By Lemma 2.25,

$$\mathbb{E}(Z_1^B\mathbb{1}_E) = \mathbb{E}(\mathbb{1}_E\mathbb{E}(Z_1^B|\mathcal{F}_t)) = \mathbb{E}(\mathbb{1}_E\boldsymbol{A}(t)e^{\boldsymbol{C}h}\boldsymbol{v}(B)), \quad \text{and}$$
$$\begin{aligned}\mathbb{E}(Z_2^B\mathbb{1}_E) &= \mathbb{E}\left(\mathbb{1}_E\frac{\boldsymbol{A}(t)e^{\boldsymbol{C}h}\boldsymbol{v}(B)}{\boldsymbol{A}(t)e^{\boldsymbol{C}h}\boldsymbol{e}}\mathbb{P}(N((t,t+h]) = 0|\mathcal{F}_t)\right)\\ &= \mathbb{E}(\mathbb{1}_E\boldsymbol{A}(t)e^{\boldsymbol{C}h}\boldsymbol{v}(B)),\end{aligned}$$

so (2.3.5) follows. This means that on the event $\{N((t,t+h]) = 0\}$,

$$\begin{aligned}\frac{\boldsymbol{A}(t+h)-\boldsymbol{A}(t)}{h} &= \frac{1}{h}\left(\frac{\boldsymbol{A}(t)e^{\boldsymbol{C}h}}{\boldsymbol{A}(t)e^{\boldsymbol{C}h}\boldsymbol{e}} - \boldsymbol{A}(t)\right)\\ &= \frac{1}{h\boldsymbol{A}(t)e^{\boldsymbol{C}h}\boldsymbol{e}}\left(\boldsymbol{A}(t)e^{\boldsymbol{C}h} - \boldsymbol{A}(t)e^{\boldsymbol{C}h}\boldsymbol{e}\cdot\boldsymbol{A}(t)\right)\\ &= \frac{1}{h\boldsymbol{A}(t)e^{\boldsymbol{C}h}\boldsymbol{e}}\{\boldsymbol{A}(t)(\boldsymbol{I}++\boldsymbol{C}h+o(h))\\ &\qquad\qquad - \boldsymbol{A}(t)(\boldsymbol{I}+\boldsymbol{C}h+o(h))\boldsymbol{e}\cdot\boldsymbol{A}(t)\}\\ &= \frac{1}{\boldsymbol{A}(t)e^{\boldsymbol{C}h}\boldsymbol{e}}\left(\boldsymbol{A}(t)\boldsymbol{C} - \boldsymbol{A}(t)\boldsymbol{C}\boldsymbol{e}\cdot\boldsymbol{A}(t)+o(1)\right),\end{aligned}$$

where the last equality follows from $\boldsymbol{A}(t)\boldsymbol{e} = 1$. Thus, if no arrival happens in some neighborhood of t, we have that

$$\boldsymbol{A}'(t) = \boldsymbol{A}(t)\boldsymbol{C} - \boldsymbol{A}(t)\boldsymbol{C}\boldsymbol{e}\cdot\boldsymbol{A}(t).$$

Now, let us prove that

$$\mathbb{1}_{N((t,t+h])>0}\boldsymbol{A}(t+h) = \mathbb{1}_{N((t,t+h])>0}\frac{\boldsymbol{A}(t)\boldsymbol{D}}{\boldsymbol{A}(t)\boldsymbol{D}\boldsymbol{e}} + o(h) \quad \text{a.s..} \tag{2.3.7}$$

The linear independence of $\boldsymbol{v}$ implies that to prove (2.3.7) is equivalent to prove that $Z_3^B = Z_4^B + o(h)$ a.s. for all $B \in \mathbb{B}(\mathscr{N})$, where

$$Z_3^B := \mathbb{P}(\theta_{t+h}N \in B, N((t,t+h]) > 0|\mathcal{F}_{t+h}), \quad \text{and}$$
$$Z_4^B := \mathbb{1}_{N((t,t+h])>0}\frac{\boldsymbol{A}(t)\boldsymbol{D}\boldsymbol{v}(B)}{\boldsymbol{A}(t)\boldsymbol{D}\boldsymbol{e}}.$$

Fix $B \in \mathbb{B}(\mathscr{N})$. As in (2.3.6), it suffices to prove that

$$\mathbb{E}(Z_3^B\mathbb{1}_E\mathbb{1}_F) = \mathbb{E}(Z_4^B\mathbb{1}_E\mathbb{1}_F) + o(h), \tag{2.3.8}$$

for all $E \in \mathcal{F}_t$ and $F \in \sigma(\{N_s\}_{s=t}^{t+h})$. In fact, by Dynkin's $\pi - -\lambda$ theorem it is enough to verify (2.3.8) with sets of the type $F_{h'} := \{N((t,t+h'] > 0)\}$ $(h' \in (0,h])$ instead of $F \in \sigma(\{N_s\}_{s=t}^{t+h})$. Hence,

$$\mathbb{E}(Z_3^B\mathbb{1}_E\mathbb{1}_{F_{h'}}) = \mathbb{E}(\mathbb{1}_E\mathbb{E}(Z_3^B\mathbb{1}_{F_{h'}}|\mathcal{F}_t)) = \mathbb{E}(\mathbb{1}_E\boldsymbol{A}(t)\boldsymbol{D}\boldsymbol{v}(B)h') + o(h'), \quad \text{and}$$
$$\begin{aligned}\mathbb{E}(Z_4^B\mathbb{1}_E\mathbb{1}_{F_{h'}}) &= \mathbb{E}\left(\mathbb{1}_E\frac{\boldsymbol{A}(t)\boldsymbol{D}\boldsymbol{v}(B)}{\boldsymbol{A}(t)\boldsymbol{D}\boldsymbol{e}}\mathbb{P}(N((t,t+h]) > 0, F_{h'}|\mathcal{F}_t)\right)\\ &= \mathbb{E}(\mathbb{1}_E\boldsymbol{A}(t)\boldsymbol{D}\boldsymbol{v}(B)h'),\end{aligned}$$

so that (2.3.8) and (2.3.7) follow. This means that if T denotes an arrival epoch, then

$$\boldsymbol{A}(T) = \frac{\boldsymbol{A}(T^-)\boldsymbol{D}}{\boldsymbol{A}(T^-)\boldsymbol{D}\boldsymbol{e}}.$$

Sequential concatenation of the previous at the arrival epochs $\{T_i\}_{i\geq 1}$ yields (2.3.4). □

Proposition 2.27 implies that we can think of $\{\boldsymbol{A}(t)\}_{t\geq 0}$ as a piecewise deterministic Markov process with state–space $\mathfrak{B}$ and local characteristics $(\mathfrak{X}, \lambda, Q)$ given by

$$\mathfrak{X}f(\boldsymbol{a}) = \sum_{i=1}^{p}\frac{\partial f}{\partial a^i}(\boldsymbol{a})\cdot((\boldsymbol{a}\boldsymbol{C})_i + \boldsymbol{a}\boldsymbol{C}\boldsymbol{e}\cdot a_i)$$
$$\lambda(\boldsymbol{a}) = \boldsymbol{a}\boldsymbol{D}\boldsymbol{e}, \quad Q\left(\boldsymbol{a}; \left\{\frac{\boldsymbol{a}\boldsymbol{D}}{\boldsymbol{a}\boldsymbol{D}\boldsymbol{e}}\right\}\right) = 1.$$

In particular, $\{\boldsymbol{A}(t)\}_{t\geq 0}$ is a strong Markov process. We call $\{\boldsymbol{A}(t)\}_{t\geq 0}$ the **associated orbit process** of parameters $(\boldsymbol{C}, \boldsymbol{D})$. The following is an alternative representation of the orbit process.

Theorem 2.28 *If $T_0 = 0$ and $\{T_i\}_{i\geq 0}$ are the arrival epochs of N, then*

$$\boldsymbol{A}(t) = \frac{\boldsymbol{A}(0)\left(\prod_{i=1}^{N_t} e^{\boldsymbol{C}(T_i - T_{i-1})}\boldsymbol{D}\right) e^{\boldsymbol{C}(t-T_{N_t})}}{\boldsymbol{A}(0)\left(\prod_{i=1}^{N_t} e^{\boldsymbol{C}(T_i - T_{i-1})}\boldsymbol{D}\right) e^{\boldsymbol{C}(t-T_{N_t})}\boldsymbol{e}}. \tag{2.3.9}$$

Proof. Let us prove (2.3.9) by induction. If $t \in (0, T_1)$, then (2.3.9) follows by (2.3.5). If $t = T_1$, then

$$\boldsymbol{A}(t) = \frac{\boldsymbol{A}(t^-)\boldsymbol{D}}{\boldsymbol{A}(t^-)\boldsymbol{D}\boldsymbol{e}} = \frac{\boldsymbol{A}(0)e^{\boldsymbol{C}t}\boldsymbol{D}}{\boldsymbol{A}(0)e^{\boldsymbol{C}t}\boldsymbol{D}\boldsymbol{e}},$$

so that (2.3.9) follows.

Now, suppose that for fixed $n \geq 2$, (2.3.9) is valid for $t \in [0, T_n]$. Then, for any $t \in (T_n, T_{n+1})$,

$$\begin{aligned}
\boldsymbol{A}(t) &= \frac{\boldsymbol{A}(T_n)e^{\boldsymbol{C}(t-T_n)}}{\boldsymbol{A}(T_n)e^{\boldsymbol{C}(t-T_n)}\boldsymbol{e}} \\
&= \frac{\left\{\frac{\boldsymbol{A}(0)(\prod_{i=1}^{n} e^{\boldsymbol{C}(T_i-T_{i-1})}\boldsymbol{D})}{\boldsymbol{A}(0)(\prod_{i=1}^{n} e^{\boldsymbol{C}(T_i-T_{i-1})}\boldsymbol{D})\boldsymbol{e}}\right\} e^{\boldsymbol{C}(t-T_n)}}{\left\{\frac{\boldsymbol{A}(0)(\prod_{i=1}^{n} e^{\boldsymbol{C}(T_i-T_{i-1})}\boldsymbol{D})}{\boldsymbol{A}(0)(\prod_{i=1}^{n} e^{\boldsymbol{C}(T_i-T_{i-1})}\boldsymbol{D})\boldsymbol{e}}\right\} e^{\boldsymbol{C}(t-T_n)}\boldsymbol{e}} \\
&= \frac{\boldsymbol{A}(0)\left(\prod_{i=1}^{n} e^{\boldsymbol{C}(T_i-T_{i-1})}\boldsymbol{D}\right) e^{\boldsymbol{C}(t-T_n)}}{\boldsymbol{A}(0)\left(\prod_{i=1}^{n} e^{\boldsymbol{C}(T_i-T_{i-1})}\boldsymbol{D}\right) e^{\boldsymbol{C}(t-T_n)}\boldsymbol{e}}.
\end{aligned}$$

If $t = T_{n+1}$, then

$$\begin{aligned}
\boldsymbol{A}(t) &= \frac{\boldsymbol{A}(t^-)\boldsymbol{D}}{\boldsymbol{A}(t^-)\boldsymbol{D}\boldsymbol{e}} \\
&= \frac{\left\{\frac{\boldsymbol{A}(0)(\prod_{i=1}^{n} e^{\boldsymbol{C}(T_i-T_{i-1})}\boldsymbol{D})e^{\boldsymbol{C}(t-T_n)}}{\boldsymbol{A}(0)(\prod_{i=1}^{n} e^{\boldsymbol{C}(T_i-T_{i-1})}\boldsymbol{D})e^{\boldsymbol{C}(t-T_n)}\boldsymbol{e}}\right\}\boldsymbol{D}}{\left\{\frac{\boldsymbol{A}(0)(\prod_{i=1}^{n} e^{\boldsymbol{C}(T_i-T_{i-1})}\boldsymbol{D})e^{\boldsymbol{C}(t-T_n)}}{\boldsymbol{A}(0)(\prod_{i=1}^{n} e^{\boldsymbol{C}(T_i-T_{i-1})}\boldsymbol{D})e^{\boldsymbol{C}(t-T_n)}\boldsymbol{e}}\right\}\boldsymbol{D}\boldsymbol{e}} \\
&= \frac{\boldsymbol{A}(0)\left(\prod_{i=1}^{n} e^{\boldsymbol{C}(T_i-T_{i-1})}\boldsymbol{D}\right) e^{\boldsymbol{C}(t-T_n)}\boldsymbol{D}}{\boldsymbol{A}(0)\left(\prod_{i=1}^{n} e^{\boldsymbol{C}(T_i-T_{i-1})}\boldsymbol{D}\right) e^{\boldsymbol{C}(t-T_n)}\boldsymbol{D}\boldsymbol{e}} \\
&= \frac{\boldsymbol{A}(0)\left(\prod_{i=1}^{n+1} e^{\boldsymbol{C}(T_i-T_{i-1})}\boldsymbol{D}\right)}{\boldsymbol{A}(0)\left(\prod_{i=1}^{n+1} e^{\boldsymbol{C}(T_i-T_{i-1})}\boldsymbol{D}\right)\boldsymbol{e}},
\end{aligned}$$

so that (2.3.9) follows for $t \in [0, T_{n+1}]$ and the proof is finished. □

The following provides an alternative characterisation of the Rational arrival process.

Theorem 2.29 *An arrival process N is a RAP if and only if there exist matrices $\boldsymbol{C}$ and $\boldsymbol{D}$, and a row vector $\boldsymbol{\alpha}$ such that* $\mathrm{dev}(\boldsymbol{C}) < 0$, $\mathrm{dev}(\boldsymbol{C}+\boldsymbol{D}) = 0$, $(\boldsymbol{C}+\boldsymbol{D})\boldsymbol{e} = 0$ *and for all* $n \geq 1$,

$$\begin{aligned}&\mathbb{P}(T_1 \in \mathrm{d}x_1, T_2 - T_1 \in \mathrm{d}x_2, \ldots, T_n - T_{n-1} \in \mathrm{d}x_n)\\ &\quad = \boldsymbol{\alpha} e^{\boldsymbol{C}x_1}\boldsymbol{D}e^{\boldsymbol{C}x_2}\boldsymbol{D}\ldots e^{\boldsymbol{C}x_n}\boldsymbol{D}\boldsymbol{e}\mathrm{d}x_1\cdots\mathrm{d}x_n, \quad x_1,\ldots,x_n \geq 0, \end{aligned} \tag{2.3.10}$$

where $T_1, T_2, \ldots$ correspond to the arrival epochs of N.

Proof. Suppose that N is a RAP. Let $\{\boldsymbol{A}(t)\}_{t\geq 0}$ be the associated orbit process of parameters $(\boldsymbol{C},\boldsymbol{D})$, and define $\boldsymbol{\alpha} = \boldsymbol{A}(0)$. According to Lemma 2.25,

$$\mathbb{P}(T_1 > x_1) = \boldsymbol{\alpha}e^{\boldsymbol{C}x_1}\boldsymbol{e}, \quad x_1 \geq 0,$$

so that

$$\mathbb{P}(T_1 \in \mathrm{d}x_1) = \boldsymbol{\alpha}e^{\boldsymbol{C}x_1}\boldsymbol{C}\boldsymbol{e}\mathrm{d}x_1 = \boldsymbol{\alpha}e^{\boldsymbol{C}x_1}\boldsymbol{D}\boldsymbol{e}\mathrm{d}x_1,$$

where the last equality follows from $(\boldsymbol{C}+\boldsymbol{D})\boldsymbol{e} = \boldsymbol{0}$. Thus, (2.3.10) follows for the case $n = 1$.

Now, suppose that (2.3.10) is true for some $n \geq 1$. Then,

$$\begin{aligned}&\mathbb{P}(T_1 \in \mathrm{d}x_1, \ldots, T_n - T_{n-1} \in \mathrm{d}x_n, T_{n+1} - T_n > x_{n+1})\\ &\quad = \mathbb{E}(\mathbb{1}\{T_1 \in \mathrm{d}x_1, \ldots, T_n - T_{n-1} \in \mathrm{d}x_n\}\mathbb{P}(T_{n+1} - T_n > x_{n+1} \mid T_1,\ldots,T_n))\\ &\quad = \mathbb{E}\left(\mathbb{1}\{T_1 \in \mathrm{d}x_1, \ldots, T_n - T_{n-1} \in \mathrm{d}x_n\}(\boldsymbol{A}(T_n)e^{\boldsymbol{C}x_{n+1}}\boldsymbol{e})\right)\\ &\quad = \mathbb{E}\left(\mathbb{1}\{T_1 \in \mathrm{d}x_1, \ldots, T_n - T_{n-1} \in \mathrm{d}x_n\}\left(\frac{\boldsymbol{\alpha}\left(\prod_{i=1}^n e^{\boldsymbol{C}x_i}\boldsymbol{D}\right)}{\boldsymbol{\alpha}\left(\prod_{i=1}^n e^{\boldsymbol{C}x_i}\boldsymbol{D}\right)\boldsymbol{e}}e^{\boldsymbol{C}x_{n+1}}\boldsymbol{e}\right)\right)\\ &\quad = \left(\frac{\boldsymbol{\alpha}\left(\prod_{i=1}^n e^{\boldsymbol{C}x_i}\boldsymbol{D}\right)}{\boldsymbol{\alpha}\left(\prod_{i=1}^n e^{\boldsymbol{C}x_i}\boldsymbol{D}\right)\boldsymbol{e}}e^{\boldsymbol{C}x_{n+1}}\boldsymbol{e}\right)\mathbb{P}(T_1 \in \mathrm{d}x_1, \ldots, T_n - T_{n-1} \in \mathrm{d}x_n)\\ &\quad = \left(\frac{\boldsymbol{\alpha}\left(\prod_{i=1}^n e^{\boldsymbol{C}x_i}\boldsymbol{D}\right)}{\boldsymbol{\alpha}\left(\prod_{i=1}^n e^{\boldsymbol{C}x_i}\boldsymbol{D}\right)\boldsymbol{e}}e^{\boldsymbol{C}x_{n+1}}\boldsymbol{e}\right)\boldsymbol{\alpha}e^{\boldsymbol{C}x_1}\boldsymbol{D}e^{\boldsymbol{C}x_2}\boldsymbol{D}\ldots e^{\boldsymbol{C}x_n}\boldsymbol{D}\boldsymbol{e}\mathrm{d}x_1\cdots\mathrm{d}x_n\\ &\quad = \boldsymbol{\alpha}e^{\boldsymbol{C}x_1}\boldsymbol{D}e^{\boldsymbol{C}x_2}\boldsymbol{D}\ldots e^{\boldsymbol{C}x_n}\boldsymbol{D}e^{\boldsymbol{C}x_{n+1}}\boldsymbol{e}\mathrm{d}x_1\cdots\mathrm{d}x_n,\end{aligned}$$

so that

$$\begin{aligned}&\mathbb{P}(T_1 \in \mathrm{d}x_1, \ldots, T_n - T_{n-1} \in \mathrm{d}x_n, T_{n+1} - T_n \in \mathrm{d}x_{n+1})\\ &\quad = -\boldsymbol{\alpha}e^{\boldsymbol{C}x_1}\boldsymbol{D}e^{\boldsymbol{C}x_2}\boldsymbol{D}\ldots e^{\boldsymbol{C}x_n}\boldsymbol{D}e^{\boldsymbol{C}x_{n+1}}\boldsymbol{C}\boldsymbol{e}\mathrm{d}x_1\cdots\mathrm{d}x_n\mathrm{d}x_{n+1}\\ &\quad = \boldsymbol{\alpha}e^{\boldsymbol{C}x_1}\boldsymbol{D}e^{\boldsymbol{C}x_2}\boldsymbol{D}\ldots e^{\boldsymbol{C}x_n}\boldsymbol{D}e^{\boldsymbol{C}x_{n+1}}\boldsymbol{D}\boldsymbol{e}\mathrm{d}x_1\cdots\mathrm{d}x_n\mathrm{d}x_{n+1},\end{aligned}$$

and thus (2.3.10) follows for all $n \geq 1$ by induction.

Now, suppose that there exist $\boldsymbol{\alpha}$, $\boldsymbol{C}$ and $\boldsymbol{D}$ with $\mathrm{dev}(\boldsymbol{C}+\boldsymbol{D}) = 0$, $(\boldsymbol{C}+\boldsymbol{D})\boldsymbol{e} = 0$ such that (2.3.10) holds for $n \geq 1$. Then, for $n \geq 1$ and $x_1, \ldots, x_n \geq 0$,

$$\begin{aligned}&\mathbb{P}(\theta_t T_1 \in \mathrm{d}x_1, \theta_t T_2 - \theta_t T_1 \in \mathrm{d}x_2, \ldots, \theta_t T_n - \theta_t T_{n-1} \in \mathrm{d}x_n \mid \mathcal{F}_t)\\ &\quad = \boldsymbol{\alpha}(t) e^{\boldsymbol{C}x_1} \boldsymbol{D} e^{\boldsymbol{C}x_2} \boldsymbol{D} \ldots e^{\boldsymbol{C}x_n} \boldsymbol{D} e^{\boldsymbol{C}x_{n+1}} \boldsymbol{D}\boldsymbol{e}\mathrm{d}x_1 \cdots \mathrm{d}x_n \mathrm{d}x_{n+1} \quad \text{for all} \quad t \geq 0,\end{aligned}$$

where

$$\boldsymbol{\alpha}(t) := \frac{\boldsymbol{\alpha}\left(\prod_{i=1}^{N_t} e^{\boldsymbol{C}(T_i - T_{i-1})}\boldsymbol{D}\right) e^{\boldsymbol{C}(t-T_{N_t})}}{\boldsymbol{\alpha}\left(\prod_{i=1}^{N_t} e^{\boldsymbol{C}(T_i - T_{i-1})}\boldsymbol{D}\right) e^{\boldsymbol{C}(t-T_{N_t})}\boldsymbol{e}},$$

which is $\mathcal{F}_t$–measurable. This implies that $\mathbb{P}(\theta_t N \in \cdot \mid \mathcal{F}_t)$ is finitely generated. The fact that $\mathrm{dev}(\boldsymbol{C}) < 0$ guarantees that $N((0,\infty)) = \infty$ and thus, N is a Rational arrival process. □

Example 2.4 (Markovian arrival process) *Let N be a $\mathrm{MAP}(\boldsymbol{\alpha}, \boldsymbol{C}, \boldsymbol{D})$. Then Theorem 2.18 and Theorem 2.29 imply that N is also a RAP. This proves that the Rational arrival process is indeed an extension of the Markovian arrival process, however, the probabilistic interpretation of N as a RAP is completely different from the one of a MAP. Indeed, the underlying process of the MAP is a Markov jump process, while the underlying process of a RAP is a piecewise deterministic Markov process with a noncountable state–space.*

Example 2.5 (ME–renewal process) *Consider a minimal standard matrix–exponential distribution of parameters $(\boldsymbol{\alpha}, \boldsymbol{S})$. Then, the Rational arrival process N with underlying orbit process $\{\boldsymbol{A}(t)\}_{t\geq 0}$ of parameters $\boldsymbol{C} = \boldsymbol{S}$, $\boldsymbol{D} = \boldsymbol{s}\boldsymbol{\alpha}$ and $\boldsymbol{A}(0) = \boldsymbol{\alpha}$ corresponds to a renewal process with $ME(\boldsymbol{\alpha}, \boldsymbol{S})$–distributed interarrival times. Indeed,*

$$\begin{aligned}&\mathbb{P}(T_1 \in \mathrm{d}x_1, T_2 - T_1 \in \mathrm{d}x_2, \ldots, T_n - T_{n-1} \in \mathrm{d}x_n)\\ &\quad = \boldsymbol{\alpha} e^{\boldsymbol{C}x_1} \boldsymbol{D} e^{\boldsymbol{C}x_2} \boldsymbol{D} \ldots e^{\boldsymbol{C}x_n} \boldsymbol{D}\boldsymbol{e}\mathrm{d}x_1 \cdots \mathrm{d}x_n\\ &\quad = \boldsymbol{\alpha} e^{\boldsymbol{S}x_1} \boldsymbol{s}\boldsymbol{\alpha} e^{\boldsymbol{S}x_2} (\boldsymbol{s}\boldsymbol{\alpha}) \ldots e^{\boldsymbol{S}x_n} (\boldsymbol{s}\boldsymbol{\alpha})\boldsymbol{e}\mathrm{d}x_1 \cdots \mathrm{d}x_n\\ &\quad = \prod_{i=1}^{n} (\boldsymbol{\alpha} e^{\boldsymbol{S}x_1}\boldsymbol{s})\mathrm{d}x_1 \cdots \mathrm{d}x_n.\end{aligned}$$

Moreover, (2.2.15) implies that the state–space $\mathfrak{B}$ of $\{\boldsymbol{A}(t)\}_{t\geq 0}$ is indeed minimal. If $\mathrm{ME}(\boldsymbol{\alpha}, \boldsymbol{S})$ is a matrix–exponential distribution which is not phase–type, then N is a point process which is a RAP but not a MAP.

Example 2.6 *[RAP which is not MAP nor ME–renewal process.] Just as in the case of matrix–exponential distributions, finding nontrivial examples of RAPs is*

difficult. For instance, one need to check that the orbit stays confined within some compact set, and that its jump intensities are nonnegative. The following is an example constructed in Asmussen and Bladt (1999). Let $\mathfrak{B}_0 := \{(1-2a, a, a) : a \in [-0.95, -0.74]\}$. *Then the orbit process* $\{\boldsymbol{A}(t)\}_{t\geq 0}$ *with* $\boldsymbol{A}(0) \in \mathfrak{B}_0$ *and parameters* $(\boldsymbol{C}, \boldsymbol{D})$ *with*

$$\boldsymbol{C} = \begin{pmatrix} -1 & 0 & 0 \\ -2/3 & -1 & 1 \\ 2/3 & -1 & -1 \end{pmatrix} \qquad \textit{and} \qquad \boldsymbol{D} = \begin{pmatrix} 14/5 & -9/10 & -9/10 \\ 26/15 & -8/15 & -8/15 \\ 58/15 & -19/15 & -19/15 \end{pmatrix}$$

constitutes the underlying process of a RAP. Furthermore, the jumps of $\{\boldsymbol{A}(t)\}_{t\geq 0}$ *always land in* $\mathfrak{B}_0 \subseteq \mathfrak{B}$.

Remark 6 *Mitchell (2001) proposes a way to construct a sequence of correlated random variables* $\{X_i\}_{i\geq 0}$ *with a common matrix–exponential distribution* $(\boldsymbol{\alpha}, \boldsymbol{S})$ *by defining*

$$\begin{aligned} &\mathbb{P}(X_1 \in \mathrm{d}x_1, X_2 \in \mathrm{d}x_2, \ldots, X_n \in \mathrm{d}x_n) \\ &\quad = \boldsymbol{\alpha} e^{\boldsymbol{S}x_1} \boldsymbol{L}_\gamma e^{\boldsymbol{S}x_1} \boldsymbol{L}_\gamma \cdots e^{\boldsymbol{S}x_2} \boldsymbol{s} \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_n, \end{aligned} \tag{2.3.11}$$

where $\boldsymbol{L}_\gamma$ *is on the form*

$$\boldsymbol{L}_\gamma = (1-\gamma)(\boldsymbol{S}\boldsymbol{e}'\boldsymbol{\alpha} - \boldsymbol{S}) + \boldsymbol{S}, \quad \gamma \in [-1, 1).$$

At first sight, their method provides a way to easily construct nontrivial examples of a RAP. In reality, the same restrictions explained in this subsection need to be verified for the parameters of the construction of Mitchell (2001). Such restrictions translate into verifing that (2.3.11) is nonnegative for all $n \geq 1$ *and* $x_1, \ldots, x_n \geq 0$*: this is a nontrivial matter.*

For further literature on the theory of Rational arrival processes, see Bladt and Neuts (2003), where an alternative characterisatiion via fluid containers is developed. Other characteristics of the RAP have been analyised in Telek (2011) (moment characterisation of a RAP), Buchholz and Telek (2012) (RAPs with marks) and Buchholz and Telek (2011) (representations of different sizes for a RAP). We will come back to more applications of the RAP in stochastic modelling in Chapter 7.

2.4 Fluid flow processes with or without Brownian components

Applications of what nowadays is known as a fluid flow process go back at least to Loynes (1961) where it was used to model dams and reserviors, and to Anick

et al. (1983) where it was used to model buffers in telecommunication systems. More theoretical–driven research was later carried on simultaneously by Rogers (1994), Asmussen (1995a) and Karandikar and Kulkarni (1995). In recent years, the theory of fluid flow processes has been considerably benefited from related applied areas, such as quasi–birth-death processes (see Ramaswami (1999) and Da Silva Soares and Latouche (2002)) and computational algorithms (see Bean et al. (2005), Nguyen and Poloni (2014) and Bean et al. (2018)). An updated survey on the literature of fluid flow processes can be found in Latouche and Nguyen (2018).

In few words, a fluid flow process consists of two processes evolving simultaneously: one is a Markov jump process and the other is a process which continually collects rewards from the holding times of the Markov jump process. A precise definition is the following. Let $\{J_t\}_{t\geq 0}$ denote a Markov jump process over a finite state–space $\mathfrak{E}$ partitioned into $\mathfrak{E}_\sigma \cup \mathfrak{E}_+ \cup \mathfrak{E}_- \cup \mathfrak{E}_0$, so that the infinitesimal matrix for $\{J_t\}_{t\geq 0}$ is given by

$$\mathbf{\Lambda} = \begin{pmatrix} \mathbf{\Lambda}_{\sigma\sigma} & \mathbf{\Lambda}_{\sigma+} & \mathbf{\Lambda}_{\sigma-} & \mathbf{\Lambda}_{\sigma 0} \\ \mathbf{\Lambda}_{+\sigma} & \mathbf{\Lambda}_{++} & \mathbf{\Lambda}_{+-} & \mathbf{\Lambda}_{+0} \\ \mathbf{\Lambda}_{-\sigma} & \mathbf{\Lambda}_{-+} & \mathbf{\Lambda}_{--} & \mathbf{\Lambda}_{-0} \\ \mathbf{\Lambda}_{0\sigma} & \mathbf{\Lambda}_{0+} & \mathbf{\Lambda}_{0-} & \mathbf{\Lambda}_{00} \end{pmatrix}.$$

For each $i \in \mathfrak{E}_\sigma$ let $\sigma_i \in (0,\infty)$ and $r_i \in \mathbb{R}$, for each $i \in \mathfrak{E}_+$ let $\sigma_i = 0$ and $r_i \in (0,\infty)$, for each $i \in \mathfrak{E}_-$ let $\sigma_i = 0$ and $r_i \in (-\infty, 0)$, and for each $i \in \mathfrak{E}_0$ let $\sigma_i = 0$ and $r_i = 0$. We call $\{r_i\}$ the **reward rates** and $\{\sigma_i\}$ the **Brownian noise rate**. A **fluid flow process** initiated at level $u \in \mathbb{R}$ is the Markov additive process (see Chapter X in Asmussen (2003)) $\{(V_t, J_t)\}_{t\geq 0}$ with

$$V_t = u + \int_0^t r_{J_s} \mathrm{d}s + \int_0^t \sigma_{J_s} \mathrm{d}B_s,$$

where $\{B_t\}_{t\geq 0}$ is a standard Brownian motion independent of $\{J_t\}_{t\geq 0}$. We call $\mathfrak{E}_\sigma \cup \mathfrak{E}_+ \cup \mathfrak{E}_- \cup \mathfrak{E}_0$ the **phase–space** of the fluid flow process $\{(V_t, J_t)\}_{t\geq 0}$. See Figure 2.7 for an example of the path of a fluid flow process. For sake of simplicity, throughout this subsection we suppose that $V_0 = 0$: the general case can be recovered by noting that conditional on $V_0 = 0$, then $\{(V_t + u, J_t)\}_{t\geq 0}$ corresponds to a fluid flow process initiated at level u.

2.4.1 Case $\mathfrak{E}_0 = \emptyset$

In this subsection we focus in the case the fluid flow process $\{(V_t, J_t)\}_{t\geq 0}$ is such that $\mathfrak{E}_0 = \emptyset$. In the following we study the first passage probabilities of $\{V_t\}_{t\geq 0}$.

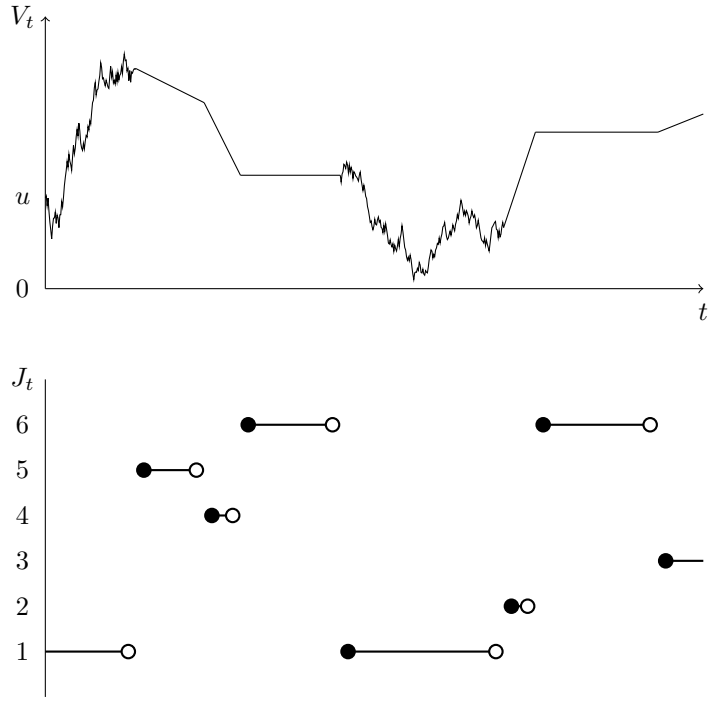


Figure 2.7: An example of a fluid flow process $\{(V_t, J_t)\}_{t\geq 0}$ started at $u > 0$ with phase–space $\mathfrak{E}_\sigma \cup \mathfrak{E}_+ \cup \mathfrak{E}_- \cup \mathfrak{E}_0$ with $\mathfrak{E}_\sigma = \{1\}$, $\mathfrak{E}_+ = \{2, 3\}$, $\mathfrak{E}_- = \{4, 5\}$ and $\mathfrak{E}_0 = \{6\}$.

2.4.1.1 Downcrossing probabilities

We are interested in computing the probability that $\{V_t\}_{t\geq 0}$ downcrosses 0 or any other arbitrary but fixed level below 0. To do this, we need to recall the Wiener–Hopf factorisation for Brownian motion.

Lemma 2.30 (Wiener–Hopf factorisation for Brownian motion) *Let $\{B_t\}_{t\geq 0}$ be a Brownian motion with variance $\sigma^2 > 0$ and drift r. Let T be an exponential random variable of rate λ, independent of $\{B_t\}_{t\geq 0}$. Then $-\min_{0\leq t\leq T} B_t$ and $B_T - \min_{0\leq t\leq T} B_t$ are independent and exponentially distributed with rates*

$$\omega = \frac{r}{\sigma^2} + \sqrt{\frac{r^2}{\sigma^4} + \frac{2\lambda}{\sigma^2}} \quad \textit{and} \quad \eta = -\frac{r}{\sigma^2} + \sqrt{\frac{r^2}{\sigma^4} + \frac{2\lambda}{\sigma^2}}, \quad \textit{respectively.}$$

In the following, let

$$\begin{aligned}
\boldsymbol{\Delta}_{r_\sigma} &= \operatorname{diag}\{r_i : i \in \mathfrak{E}_\sigma\},\\
\boldsymbol{\Delta}_{r_+} &= \operatorname{diag}\{r_i : i \in \mathfrak{E}_+\},\\
\boldsymbol{\Delta}_{r_-} &= \operatorname{diag}\{r_i : i \in \mathfrak{E}_-\},\\
\boldsymbol{\Delta}_{\sigma} &= \operatorname{diag}\{\sigma_i : i \in \mathfrak{E}_\sigma\},\\
\boldsymbol{\Delta}_{\lambda^\sigma} &= \operatorname{diag}\{-\lambda_{ii} : i \in \mathfrak{E}_\sigma\},\\
\boldsymbol{\Delta}_{\omega} &= \operatorname{diag}\left\{\frac{r_i}{\sigma_i^2} + \sqrt{\frac{r_i^2}{\sigma_i^4} + \frac{-2\lambda_{ii}}{\sigma_i^2}} : i \in \mathfrak{E}_\sigma\right\}, \qquad (2.4.1)\\
\boldsymbol{\Delta}_{\eta} &= \operatorname{diag}\left\{-\frac{r_i}{\sigma_i^2} + \sqrt{\frac{r_i^2}{\sigma_i^4} + \frac{-2\lambda_{ii}}{\sigma_i^2}} : i \in \mathfrak{E}_\sigma\right\}. \qquad (2.4.2)
\end{aligned}$$

We are interested in computing first passage probabilities of $\{(V_t, J_t)\}_{t\geq 0}$, that is, the distribution of J_τ and $m_x = J_{\tau_x}$, where $\tau := \inf\{t > 0 : V_t < 0\}$ and $\tau_x = \inf\{t > 0 : V_t < -x\}$. One can argue that, by the strong Markov property of $\{J_t\}_{t\geq 0}$, the process $\{m_x\}_{x\geq 0}$ is a possibly terminating Markov jump process with state–space $\mathfrak{E}_\sigma \cup \mathfrak{E}_-$ and, say, infinitesimal matrix $\boldsymbol{D}$. Furthermore, for $i \in \mathfrak{E}_+$ let

$$\begin{aligned}
\beta_{ij}^{+\sigma} &= \mathbb{P}(J_\tau = j \mid J_0 = i, V_0 = 0), \quad j \in \mathfrak{E}_\sigma,\\
\beta_{ij}^{+-} &= \mathbb{P}(J_\tau = j \mid J_0 = i, V_0 = 0), \quad j \in \mathfrak{E}_-,
\end{aligned}$$

and define the matrices $\boldsymbol{\beta}_{+\sigma} = \{\beta_{ij}^{+\sigma}\}_{i\in\mathfrak{E}_+, j\in\mathfrak{E}_\sigma}$ and $\boldsymbol{\beta}_{+-} = \{\beta_{ij}^{+-}\}_{i\in\mathfrak{E}_+, j\in\mathfrak{E}_-}$. To ease notation, let

$$\boldsymbol{\Lambda}_{\sigma d} = \begin{pmatrix}\boldsymbol{\Lambda}_{\sigma\sigma} & \boldsymbol{\Lambda}_{\sigma-}\end{pmatrix}, \quad \boldsymbol{\Lambda}_{+d} = \begin{pmatrix}\boldsymbol{\Lambda}_{+\sigma} & \boldsymbol{\Lambda}_{+-}\end{pmatrix}, \quad \boldsymbol{\Lambda}_{-d} = \begin{pmatrix}\boldsymbol{\Lambda}_{-\sigma} & \boldsymbol{\Lambda}_{--}\end{pmatrix},$$

$$\boldsymbol{D} = \begin{pmatrix}\boldsymbol{D}_{\sigma\sigma} & \boldsymbol{D}_{\sigma-}\\ \boldsymbol{D}_{-\sigma} & \boldsymbol{D}_{--}\end{pmatrix} =: \begin{pmatrix}\boldsymbol{D}_{\sigma d}\\ \boldsymbol{D}_{-d}\end{pmatrix}, \quad \text{and}$$

$$\boldsymbol{\beta}_{+d} = \begin{pmatrix}\boldsymbol{\beta}_{+\sigma} & \boldsymbol{\beta}_{+-}\end{pmatrix}.$$

The following are properties of the matrices $\boldsymbol{D}$ and $\boldsymbol{\beta}_{+d}$.

Theorem 2.31 *The matrices $\boldsymbol{\beta}_{+d}$ and $\boldsymbol{D}$ satisfy the equations*

$$\begin{aligned}
&\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{+d} + \boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{++}\boldsymbol{\beta}_{+d} + \boldsymbol{\beta}_{+d}\boldsymbol{D} = \mathbf{0}, && (2.4.3)\\
&\boldsymbol{D}_{-d} = |\boldsymbol{\Delta}_{r_-}^{-1}|(\boldsymbol{\Lambda}_{-d} + \boldsymbol{\Lambda}_{-+}\boldsymbol{\beta}_{+d}), && (2.4.4)\\
&\boldsymbol{D}_{\sigma d}\boldsymbol{D} + 2\boldsymbol{\Delta}_\sigma^{-2}(\boldsymbol{\Delta}_{r_\sigma}\boldsymbol{D}_{\sigma d} + \boldsymbol{\Lambda}_{\sigma d} + \boldsymbol{\Lambda}_{\sigma+}\boldsymbol{\beta}_{+d}) = \mathbf{0}. && (2.4.5)
\end{aligned}$$

PROOF. Notice that the level at which $\{V_t\}_{t\geq 0}$ changes direcction, from going upwards to going downwards, is phase–type–distributed with sub–intensity matrix $\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{++}$: this follows by arguments similar to those used to prove (2.2.11). Now, to compute $\boldsymbol{\beta}_{+d}$ we may condition on the level $\{V_t\}_{t\geq 0}$ reaches before changing direction for the first time (which happens with intensity $\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{+d}$), point at which a downwards process governed by the sub–intensity matrix $\boldsymbol{D}$ takes place until $\{V_t\}_{t\geq 0}$ reaches 0 again. That is,

$$\boldsymbol{\beta}_{+d} = \int_0^\infty e^{\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{++}y}\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{+d}e^{\boldsymbol{D}y}\mathrm{d}y.$$

Using integration by parts, this is equivalent to

$$\begin{aligned}\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{++}\boldsymbol{\beta}_{+d} &= \int_0^\infty \boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{++}e^{\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{++}y}\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{+d}e^{\boldsymbol{D}y}\mathrm{d}y \\ &= e^{\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{++}y}\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{+d}e^{\boldsymbol{D}y}\Big|_{y=0}^{\infty} - \int_0^\infty e^{\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{++}y}\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{+d}e^{\boldsymbol{D}y}\boldsymbol{D}\mathrm{d}y \\ &= -\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{+d} - \boldsymbol{\beta}_{+d}\boldsymbol{D},\end{aligned}$$

and so (2.4.3) is proved.

Equation (2.4.4) follows by noticing that if $J_0 = i \in \mathfrak{E}_-$, then the downward process with sub–intensity matrix $\boldsymbol{D}$ can downcross level $-\mathrm{d}x$ by making one jump to $j \in \mathfrak{E}_\sigma \cup \mathfrak{E}_-$ (zero jumps if $i = j$) with probability $\delta_{ij} + |1/r_i|(\boldsymbol{\Lambda}_{-d})_{ij}\mathrm{d}x$, or by making a jump to some state in $\mathfrak{E}_+$ and eventually downcrossing level $-\mathrm{d}x$ while in state $k \in \mathfrak{E}_\sigma \cup \mathfrak{E}_-$, which happens with probability $|1/r_i|(\boldsymbol{\Lambda}_{-+}\boldsymbol{\beta}_{+d})_{ik}\mathrm{d}x$.

To prove (2.4.5), notice that

$$\begin{pmatrix}\boldsymbol{I} + \boldsymbol{D}_{\sigma\sigma} & \boldsymbol{D}_{\sigma-}\end{pmatrix}\mathrm{d}x = \begin{pmatrix}\boldsymbol{I} - \boldsymbol{\Delta}_\omega & \boldsymbol{0}\end{pmatrix}\mathrm{d}x + (\boldsymbol{\Delta}_\omega \mathrm{d}x)\int_0^\infty \boldsymbol{\Delta}_\eta e^{-\boldsymbol{\Delta}_\eta y}\boldsymbol{Q}_{\sigma d}e^{\boldsymbol{D}y}\mathrm{d}y, \tag{2.4.6}$$

where

$$\boldsymbol{Q}_{\sigma d} = [\begin{pmatrix}\boldsymbol{I} & \boldsymbol{0}\end{pmatrix} + \boldsymbol{\Delta}_{\lambda^\sigma}^{-1}\boldsymbol{\Lambda}_{\sigma d}] + \boldsymbol{\Delta}_{\lambda^\sigma}^{-1}\boldsymbol{\Lambda}_{\sigma+}\boldsymbol{\beta}_{+d}. \tag{2.4.7}$$

Indeed, if $J_0 = i \in \mathfrak{E}_\sigma$, the first summand in the right hand side of (2.4.6) corresponds to the intensity of reaching level $-\mathrm{d}x$ without having jumps (given by $1 - \omega_i \mathrm{d}x$). The second summand corresponds to the intensity of not reaching level $-\mathrm{d}x$ while in i (given by $\omega_i \mathrm{d}x$), and integrating over the level $y \in [0, \infty)$ at which the process jumps to any other state. After such a jump, the process must eventually downcross level y in some state $j \in \mathfrak{E}_\sigma \cup \mathfrak{E}_-$ (which happens with probability $(\boldsymbol{Q}_{\sigma d})_{ij}$). Following this, on a level–sense a Markov jump process with intensity matrix $\boldsymbol{D}$ takes place until it finally reaches level $-\mathrm{d}x$.

Thus,

$$\boldsymbol{D}_{\sigma d} = \begin{pmatrix} -\boldsymbol{\Delta}_\omega & \mathbf{0} \end{pmatrix} + \boldsymbol{\Delta}_\omega \int_0^\infty \boldsymbol{\Delta}_\eta e^{-\boldsymbol{\Delta}_\eta y} \boldsymbol{Q}_{\sigma d} e^{\boldsymbol{D} y} \mathrm{d}y, \tag{2.4.8}$$

which in turn implies that

$$\int_0^\infty e^{-\boldsymbol{\Delta}_\eta y} \boldsymbol{Q}_{\sigma d} e^{\boldsymbol{D} y} \mathrm{d}y = \boldsymbol{\Delta}_\omega^{-1} \boldsymbol{\Delta}_\eta^{-1} \left(\boldsymbol{D}_{\sigma d} + \begin{pmatrix} \boldsymbol{\Delta}_\omega & \mathbf{0} \end{pmatrix} \right). \tag{2.4.9}$$

Integrating by parts the right hand side of (2.4.8) and using (2.4.9) we get that

$$\begin{aligned}
\boldsymbol{D}_{\sigma d} &= \begin{pmatrix} -\boldsymbol{\Delta}_\omega & \mathbf{0} \end{pmatrix} + \boldsymbol{\Delta}_\omega \left[-e^{-\boldsymbol{\Delta}_\eta y} \boldsymbol{Q}_{\sigma d} e^{\boldsymbol{D} y} \Big|_{y=0}^\infty + \int_0^\infty e^{-\boldsymbol{\Delta}_\eta y} \boldsymbol{Q}_{\sigma d} e^{\boldsymbol{D} y} \boldsymbol{D} \mathrm{d}y \right] \\
&= \begin{pmatrix} -\boldsymbol{\Delta}_\omega & \mathbf{0} \end{pmatrix} + \boldsymbol{\Delta}_\omega \left[\boldsymbol{Q}_{\sigma d} + \int_0^\infty e^{-\boldsymbol{\Delta}_\eta y} \boldsymbol{Q}_{\sigma d} e^{\boldsymbol{D} y} \boldsymbol{D} \mathrm{d}y \right] \\
&= \begin{pmatrix} -\boldsymbol{\Delta}_\omega & \mathbf{0} \end{pmatrix} + \boldsymbol{\Delta}_\omega \left[\boldsymbol{Q}_{\sigma d} + \boldsymbol{\Delta}_\omega^{-1} \boldsymbol{\Delta}_\eta^{-1} \left(\boldsymbol{D}_{\sigma d} + \begin{pmatrix} \boldsymbol{\Delta}_\omega & \mathbf{0} \end{pmatrix} \right) \boldsymbol{D} \right] \\
&= \begin{pmatrix} -\boldsymbol{\Delta}_\omega & \mathbf{0} \end{pmatrix} + \boldsymbol{\Delta}_\omega \boldsymbol{Q}_{\sigma d} + \boldsymbol{\Delta}_\eta^{-1} \boldsymbol{D}_{\sigma d} \boldsymbol{D} + \begin{pmatrix} \boldsymbol{\Delta}_\eta^{-1} \boldsymbol{\Delta}_\omega & \mathbf{0} \end{pmatrix} \boldsymbol{D} \\
&= \begin{pmatrix} -\boldsymbol{\Delta}_\omega & \mathbf{0} \end{pmatrix} + \boldsymbol{\Delta}_\omega \boldsymbol{Q}_{\sigma d} + \boldsymbol{\Delta}_\eta^{-1} \boldsymbol{D}_{\sigma d} \boldsymbol{D} + \boldsymbol{\Delta}_\eta^{-1} \boldsymbol{\Delta}_\omega \boldsymbol{D}_{\sigma d}.
\end{aligned}$$

Premultiplying by $\boldsymbol{\Delta}_\eta$,

$$\boldsymbol{\Delta}_\eta \boldsymbol{D}_{\sigma d} = \begin{pmatrix} -\boldsymbol{\Delta}_\eta \boldsymbol{\Delta}_\omega & \mathbf{0} \end{pmatrix} + \boldsymbol{\Delta}_\eta \boldsymbol{\Delta}_\omega \boldsymbol{Q}_{\sigma d} + \boldsymbol{D}_{\sigma d} \boldsymbol{D} + \boldsymbol{\Delta}_\omega \boldsymbol{D}_{\sigma d}, \quad \text{or}$$

$$\boldsymbol{D}_{\sigma d} \boldsymbol{D} + (\boldsymbol{\Delta}_\omega - \boldsymbol{\Delta}_\eta) \boldsymbol{D}_{\sigma d} - \boldsymbol{\Delta}_\eta \boldsymbol{\Delta}_\omega \left(\begin{pmatrix} \boldsymbol{I} & \mathbf{0} \end{pmatrix} + \boldsymbol{Q}_{\sigma d} \right) = \mathbf{0} \tag{2.4.10}$$

By (2.4.1), (2.4.2) and (2.4.7),

$$\boldsymbol{\Delta}_\omega - \boldsymbol{\Delta}_\eta = 2\boldsymbol{\Delta}_\sigma^{-2} \boldsymbol{\Delta}_{r_\sigma}, \quad \boldsymbol{\Delta}_\eta \boldsymbol{\Delta}_\omega = 2\boldsymbol{\Delta}_\sigma^{-2} \boldsymbol{\Delta}_{\lambda\sigma} \quad \text{and}$$

$$\begin{pmatrix} \boldsymbol{I} & \mathbf{0} \end{pmatrix} + \boldsymbol{Q}_{\sigma d} = \boldsymbol{\Delta}_{\lambda^\sigma}^{-1} \left(\boldsymbol{\Lambda}_{\sigma d} + \boldsymbol{\Lambda}_{\sigma +} \boldsymbol{\beta}_{+d} \right).$$

We get (2.4.5) by substituting the previous in (2.4.10) and the proof is finished. □

Remark 7 *In the case the state space of* $\{J_t\}_{t \geq 0}$ *is* $\mathfrak{E}_+ \cup \mathfrak{E}_-$*, the equations in Theorem 2.31 can be simplified into*

$$\boldsymbol{\Delta}_{r_+}^{-1} \boldsymbol{\Lambda}_{++} \boldsymbol{\beta}_{+-} + \boldsymbol{\beta}_{+-} |\boldsymbol{\Delta}_{r_-}^{-1}| \boldsymbol{\Lambda}_{--} + \boldsymbol{\beta}_{+-} |\boldsymbol{\Delta}_{r_-}^{-1}| \boldsymbol{\Lambda}_{-+} \boldsymbol{\beta}_{+-} + \boldsymbol{\Delta}_{r_+}^{-1} \boldsymbol{\Lambda}_{+-} = \mathbf{0}, \quad \textit{and} \tag{2.4.11}$$

$$\boldsymbol{D} = |\boldsymbol{\Delta}_{r_-}^{-1}| (\boldsymbol{\Lambda}_{--} + \boldsymbol{\Lambda}_{-+} \boldsymbol{\beta}_{+-}). \tag{2.4.12}$$

In fact, (2.4.11) is known as a **Riccati equation** *which needs to be solved for* $\boldsymbol{\beta}_{+-}$*. Once this is done, we can determine* $\boldsymbol{D}$ *by (2.4.12). See Bean et al. (2005) for a survey on additional extra methods to solve the equation (2.4.11) associated to a fluid flow process.*

Remark 8 *In the case the state space of* $\{J_t\}_{t\geq 0}$ *is* $\mathfrak{E}_\sigma$*, the equations in Theorem 2.31 can be simplified into the single matrix equation*

$$\boldsymbol{D}^2 + 2\boldsymbol{\Delta}_\sigma^{-2}\boldsymbol{\Delta}_{r_\sigma}\boldsymbol{D} + 2\boldsymbol{\Delta}_\sigma^{-2}\boldsymbol{\Lambda} = \boldsymbol{0}, \tag{2.4.13}$$

which corresponds to a **quadratic matrix equation** *which needs to be solved for* $\boldsymbol{D}$*. See Nguyen and Poloni (2016) for an optimized algorithm to solve (2.4.13).*

The following algorithm provides an iterative method to compute $\boldsymbol{\beta}_{+d}$ and $\boldsymbol{D}$. Its proof relies on pathwise arguments and probabilistic interpretations of the equations in Theorem 2.31; for more details we refer the reader to Theorem 2.4 in Simon (2017).

Theorem 2.32 *Let*

$$\boldsymbol{D}^{(0)} = \begin{pmatrix} -\boldsymbol{\Delta}_\omega & \boldsymbol{0} \\ |\boldsymbol{\Delta}_{r_-}^{-1}|\boldsymbol{\Lambda}_{-\sigma} & |\boldsymbol{\Delta}_{r_-}^{-1}|\boldsymbol{\Lambda}_{--} \end{pmatrix}.$$

For $n \geq 0$*, recursively define* $\{\boldsymbol{\beta}_{+d}^{(n)}\}_{n\geq 1}$ *and* $\{\boldsymbol{D}^{(n)}\}_{n\geq 1}$ *where:*

- $\boldsymbol{\beta}_{+d}^{(n)}$ *is the unique solution to the Sylvester equation*

$$\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{+d} + \boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{++}\boldsymbol{X} + \boldsymbol{X}\boldsymbol{D}^{(n-1)} = \boldsymbol{0},$$

- $\boldsymbol{D}_{-d}^{(n)}$ *is defined by*

$$\boldsymbol{D}_{-d}^{(n)} = |\boldsymbol{\Delta}_{r_-}^{-1}|(\boldsymbol{\Lambda}_{-d} + \boldsymbol{\Lambda}_{-+}\boldsymbol{\beta}_{+d}^{(n)}),$$

- $\boldsymbol{D}_{\sigma d}^{(n)}$ *is the unique solution to the Sylvester equation*

$$\boldsymbol{X}\boldsymbol{D}^{(n-1)} - \boldsymbol{\Delta}_\eta\boldsymbol{X} = -\boldsymbol{\Delta}_\omega\boldsymbol{D}_{\sigma d}^{(n-1)} - 2\boldsymbol{\Delta}_\sigma^{-2}(\boldsymbol{\Lambda}_{-d} + \boldsymbol{\Lambda}_{-+}\boldsymbol{\beta}_{+d}^{(n)}),$$

- $\boldsymbol{D}^{(n)}$ *is defined by*

$$\boldsymbol{D}^{(n)} = \begin{pmatrix} \boldsymbol{D}_{\sigma d}^{(n)} \\ \boldsymbol{D}_{-d}^{(n)} \end{pmatrix}.$$

Then

$$\boldsymbol{\beta}_{+d} = \lim_{n\to\infty} \boldsymbol{\beta}_{+d}^{(n)} \quad \textit{and} \quad \boldsymbol{D} = \lim_{n\to\infty} \boldsymbol{D}^{(n)}.$$

Once $\boldsymbol{\beta_{+d}}$ and $\boldsymbol{D}$ have been computed, we can calculate the downcrossing probabilities of $\{(V_t, J_t)\}_{t\geq 0}$ as follows.

Theorem 2.33 *For $i \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+ \cup \mathfrak{E}_-$ and $V_0 = 0$,*

$$\mathbb{P}(m_x = j, \tau_x < \infty \mid J_0 = i) = \left(\boldsymbol{e}_i' \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{\beta}_{+\sigma} & \boldsymbol{\beta}_{+-} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} e^{\boldsymbol{D}x} \right)_j .$$

In particular,

$$\mathbb{P}(\tau_x < \infty \mid J_0 = i) = \boldsymbol{e}_i' \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{\beta}_{+\sigma} & \boldsymbol{\beta}_{+-} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} e^{\boldsymbol{D}x} \boldsymbol{e}.$$

2.4.1.2 Upcrossing probabilities

Now we are interested in finding upcrossing probabilities for the process $\{(V_t, J_t)\}_{t \geq 0}$. Let $\gamma := \inf\{t > 0 : V_t > 0\}$, and for any $x \geq 0$ define

$$\gamma_x = \inf\{t > 0 : V_t > x\}, \quad \text{and}$$

$$n_x = J_{\gamma_x}.$$

The study of $\{n_x\}_{x \geq 0}$ is completely analogous to the one of $\{m_x\}_{x \geq 0}$. For sake of completeness, we state its corresponding results in Theorem 2.34 and Theorem 2.35 below.

Let $\boldsymbol{U}$ denote the intensity matrix of the Markov jump process $\{n_x\}_{x \geq 0}$ with state–space $\mathfrak{E}_\sigma \cup \mathfrak{E}_+$. For $i \in \mathfrak{E}_-$ let

$$\begin{aligned} \alpha_{ij}^{-\sigma} &= \mathbb{P}(J_\tau = j \mid J_0 = i, V_0 = 0), \quad j \in \mathfrak{E}_\sigma, \\ \alpha_{ij}^{-+} &= \mathbb{P}(J_\tau = j \mid J_0 = i, V_0 = 0), \quad j \in \mathfrak{E}_+, \end{aligned}$$

and define the matrices $\boldsymbol{\alpha}_{-\sigma} = \{\alpha_{ij}^{-\sigma}\}_{i \in \mathfrak{E}_-, j \in \mathfrak{E}_\sigma}$ and $\boldsymbol{\alpha}_{-+} = \{\alpha_{ij}^{-+}\}_{i \in \mathfrak{E}_-, j \in \mathfrak{E}_+}$. Let

$$\boldsymbol{\Lambda}_{\sigma u} = \begin{pmatrix} \boldsymbol{\Lambda}_{\sigma\sigma} & \boldsymbol{\Lambda}_{\sigma+} \end{pmatrix}, \quad \boldsymbol{\Lambda}_{+u} = \begin{pmatrix} \boldsymbol{\Lambda}_{+\sigma} & \boldsymbol{\Lambda}_{++} \end{pmatrix}, \quad \boldsymbol{\Lambda}_{-u} = \begin{pmatrix} \boldsymbol{\Lambda}_{-\sigma} & \boldsymbol{\Lambda}_{-+} \end{pmatrix},$$

$$\boldsymbol{U} = \begin{pmatrix} \boldsymbol{U}_{\sigma\sigma} & \boldsymbol{U}_{\sigma+} \\ \boldsymbol{U}_{+\sigma} & \boldsymbol{U}_{++} \end{pmatrix} =: \begin{pmatrix} \boldsymbol{U}_{\sigma u} \\ \boldsymbol{U}_{+u} \end{pmatrix}, \quad \text{and}$$

$$\boldsymbol{\alpha}_{-u} = \begin{pmatrix} \boldsymbol{\alpha}_{-\sigma} & \boldsymbol{\alpha}_{-+} \end{pmatrix}.$$

Theorem 2.34 *The matrices $\boldsymbol{\alpha}_{-u}$ and $\boldsymbol{U}$ satisfy the equations*

$$
\begin{aligned}
&|\boldsymbol{\Delta}_{r_-}^{-1}|\boldsymbol{\Lambda}_{-u} + |\boldsymbol{\Delta}_{r_-}^{-1}|\boldsymbol{\Lambda}_{--}\boldsymbol{\alpha}_{-u} + \boldsymbol{\alpha}_{-u}\boldsymbol{U} = \boldsymbol{0},\\
&\boldsymbol{U}_{+u} = \boldsymbol{\Delta}_{r_+}^{-1}(\boldsymbol{\Lambda}_{+u} + \boldsymbol{\Lambda}_{+-}\boldsymbol{\alpha}_{-u}),\\
&\boldsymbol{U}_{\sigma u}\boldsymbol{U} + 2\boldsymbol{\Delta}_{\sigma}^{-2}(-\boldsymbol{\Delta}_{r_\sigma}\boldsymbol{U}_{\sigma u} + \boldsymbol{\Lambda}_{\sigma u} + \boldsymbol{\Lambda}_{\sigma -}\boldsymbol{\alpha}_{-u}) = \boldsymbol{0}.
\end{aligned}
$$

Remark 9 *In the case the state space of $\{J_t\}_{t\geq 0}$ is $\mathfrak{E}_+ \cup \mathfrak{E}_-$, the equations in Theorem 2.34 are equivalent to*

$$
|\boldsymbol{\Delta}_{r_-}^{-1}|\boldsymbol{\Lambda}_{--}\boldsymbol{\alpha}_{-+} + \boldsymbol{\alpha}_{-+}\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{++} + \boldsymbol{\alpha}_{-+}\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{+-}\boldsymbol{\alpha}_{-+} + |\boldsymbol{\Delta}_{r_-}^{-1}|\boldsymbol{\Lambda}_{-+} = \boldsymbol{0}, \quad \textit{and} \tag{2.4.14}
$$

$$
\boldsymbol{U} = \boldsymbol{\Delta}_{r_+}^{-1}(\boldsymbol{\Lambda}_{++} + \boldsymbol{\Lambda}_{+-}\boldsymbol{\alpha}_{-+}). \tag{2.4.15}
$$

Remark 10 *In the case the state space of $\{J_t\}_{t\geq 0}$ is $\mathfrak{E}_\sigma$, the equations in Theorem 2.34 are equivalent to the quadratic matrix equation*

$$
\boldsymbol{U}^2 - 2\boldsymbol{\Delta}_{\sigma}^{-2}\boldsymbol{\Delta}_{r_\sigma}\boldsymbol{U} + 2\boldsymbol{\Delta}_{\sigma}^{-2}\boldsymbol{\Lambda} = \boldsymbol{0}. \tag{2.4.16}
$$

Similarly to Theorem 2.32, the following provides an iterative algorithm to compute the matrices $\boldsymbol{\alpha}_{-u}$ and $\boldsymbol{U}$.

Theorem 2.35 *Let*

$$
\boldsymbol{U}^{(0)} = \begin{pmatrix} -\boldsymbol{\Delta}_\eta & \boldsymbol{0} \\ \boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{+\sigma} & \boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}_{++} \end{pmatrix}.
$$

For $n \geq 0$, recursively define $\{\boldsymbol{\alpha}_{-u}^{(n)}\}_{n\geq 1}$ and $\{\boldsymbol{U}^{(n)}\}_{n\geq 1}$ where:

- *$\boldsymbol{\alpha}_{-u}^{(n)}$ is the unique solution to the Sylvester equation*

$$
|\boldsymbol{\Delta}_{r_-}^{-1}|\boldsymbol{\Lambda}_{-u} + |\boldsymbol{\Delta}_{r_-}^{-1}|\boldsymbol{\Lambda}_{--}\boldsymbol{X} + \boldsymbol{X}\boldsymbol{U}^{(n-1)} = \boldsymbol{0},
$$

- *$\boldsymbol{U}_{+u}^{(n)}$ is defined by*

$$
\boldsymbol{U}_{+u}^{(n)} = \boldsymbol{\Delta}_{r_+}^{-1}(\boldsymbol{\Lambda}_{+u} + \boldsymbol{\Lambda}_{+-}\boldsymbol{\alpha}_{-u}^{(n)}),
$$

- *$\boldsymbol{U}_{\sigma u}^{(n)}$ is the unique solution to the Sylvester equation*

$$
\boldsymbol{X}\boldsymbol{U}^{(n-1)} - \boldsymbol{\Delta}_\omega\boldsymbol{X} = -\boldsymbol{\Delta}_\eta\boldsymbol{U}_{\sigma u}^{(n-1)} - 2\boldsymbol{\Delta}_{\sigma}^{-2}(\boldsymbol{\Lambda}_{+u} + \boldsymbol{\Lambda}_{+-}\boldsymbol{\alpha}_{-u}^{(n)}),
$$

- *$\boldsymbol{U}^{(n)}$ is defined by*

$$
\boldsymbol{U}^{(n)} = \begin{pmatrix} \boldsymbol{U}_{\sigma u}^{(n)} \\ \boldsymbol{U}_{+u}^{(n)} \end{pmatrix}.
$$

Then

$$\boldsymbol{\alpha}_{-u} = \lim_{n\to\infty} \boldsymbol{\alpha}_{-u}^{(n)} \quad \textit{and} \quad \boldsymbol{U} = \lim_{n\to\infty} \boldsymbol{U}^{(n)}.$$

The following corresponds to the upcrossing probabilities of $\{(V_t, J_t)\}_{t\geq 0}$.

Theorem 2.36 *For $i \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+ \cup \mathfrak{E}_-$ and $V_0 = 0$,*

$$\mathbb{P}(n_x = j, \gamma_x < \infty \mid J_0 = i) = \left(\boldsymbol{e}_i' \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \\ \boldsymbol{\alpha}_{-\sigma} & \boldsymbol{\alpha}_{-+} \end{pmatrix} e^{\boldsymbol{U}x} \right)_j .$$

In particular,

$$\mathbb{P}(\gamma_x < \infty \mid J_0 = i) = \boldsymbol{e}_i' \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \\ \boldsymbol{\alpha}_{-\sigma} & \boldsymbol{\alpha}_{-+} \end{pmatrix} e^{\boldsymbol{U}x} \boldsymbol{e}.$$

2.4.2 Case $\mathfrak{E}_0 \neq \emptyset$

Let $\{(V_t, J_t)\}_{t\geq 0}$ be a general fluid flow process, where

- $\{J_t\}_{t\geq 0}$ has state space $\mathfrak{E}_\sigma \cup \mathfrak{E}_+ \cup \mathfrak{E}_- \cup \mathfrak{E}_0$ and infinitesimal matrix

$$\boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{\Lambda}_{\sigma\sigma} & \boldsymbol{\Lambda}_{\sigma+} & \boldsymbol{\Lambda}_{\sigma-} & \boldsymbol{\Lambda}_{\sigma 0} \\ \boldsymbol{\Lambda}_{+\sigma} & \boldsymbol{\Lambda}_{++} & \boldsymbol{\Lambda}_{+-} & \boldsymbol{\Lambda}_{+0} \\ \boldsymbol{\Lambda}_{-\sigma} & \boldsymbol{\Lambda}_{-+} & \boldsymbol{\Lambda}_{--} & \boldsymbol{\Lambda}_{-0} \\ \boldsymbol{\Lambda}_{0\sigma} & \boldsymbol{\Lambda}_{0+} & \boldsymbol{\Lambda}_{0-} & \boldsymbol{\Lambda}_{00} \end{pmatrix},$$

- $\{V_t\}_{t\geq 0}$ is defined

$$V_t = \int_0^t r_{J_s} \mathrm{d}s + \int_0^t \sigma_{J_s} \mathrm{d}B_s,$$

  where for each $i \in \mathfrak{E}_\sigma$ we let $\sigma_i \in (0, \infty)$ and $r_i \in \mathbb{R}$, for each $i \in \mathfrak{E}_+$ we let $\sigma_i = 0$ and $r_i \in (0, \infty)$, for each $i \in \mathfrak{E}_-$ we let $\sigma_i = 0$ and $r_i \in (-\infty, 0)$, and for each $i \in \mathfrak{E}_0$ we let $\sigma_i = 0$ and $r_i = 0$.

Let $\{J_t^c\}_{t\geq 0}$ be the process $\{J_t\}_{t\geq 0}$ censored at its occupation times in $\mathfrak{E}_0$. That is, we obtain a path of $\{J_t^c\}_{t\geq 0}$ by deleting all the holding times of $\{J_t\}_{t\geq 0}$ while it is in $\mathfrak{E}_0$. Using the exact same arguments used to prove (2.2.12), we get that

the state space of $\{J_t^c\}_{t\geq 0}$ is $\mathfrak{E}_\sigma \cup \mathfrak{E}_+ \cup \mathfrak{E}_-$ with intensity matrix

$$\begin{aligned}\boldsymbol{\Lambda}^c &= \begin{pmatrix} \boldsymbol{\Lambda}_{\sigma\sigma} & \boldsymbol{\Lambda}_{\sigma+} & \boldsymbol{\Lambda}_{\sigma-} \\ \boldsymbol{\Lambda}_{+\sigma} & \boldsymbol{\Lambda}_{++} & \boldsymbol{\Lambda}_{+-} \\ \boldsymbol{\Lambda}_{-\sigma} & \boldsymbol{\Lambda}_{-+} & \boldsymbol{\Lambda}_{--} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\Lambda}_{\sigma 0} \\ \boldsymbol{\Lambda}_{+0} \\ \boldsymbol{\Lambda}_{-0} \end{pmatrix} (-\boldsymbol{\Lambda}_{00})^{-1} \begin{pmatrix} \boldsymbol{\Lambda}_{0\sigma} & \boldsymbol{\Lambda}_{0+} & \boldsymbol{\Lambda}_{0-} \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{\Lambda}^c_{\sigma\sigma} & \boldsymbol{\Lambda}^c_{\sigma+} & \boldsymbol{\Lambda}^c_{\sigma-} \\ \boldsymbol{\Lambda}^c_{+\sigma} & \boldsymbol{\Lambda}^c_{++} & \boldsymbol{\Lambda}^c_{+-} \\ \boldsymbol{\Lambda}^c_{-\sigma} & \boldsymbol{\Lambda}^c_{-+} & \boldsymbol{\Lambda}^c_{--} \end{pmatrix}, \end{aligned} \tag{2.4.17}$$

where

$$\begin{aligned} \boldsymbol{\Lambda}^c_{\sigma\sigma} &= \boldsymbol{\Lambda}_{\sigma\sigma} + \boldsymbol{\Lambda}_{\sigma 0}(-\boldsymbol{\Lambda}_{00})^{-1}\boldsymbol{\Lambda}_{0\sigma}, \\ \boldsymbol{\Lambda}^c_{\sigma+} &= \boldsymbol{\Lambda}_{\sigma+} + \boldsymbol{\Lambda}_{\sigma 0}(-\boldsymbol{\Lambda}_{00})^{-1}\boldsymbol{\Lambda}_{0+}, \\ \boldsymbol{\Lambda}^c_{\sigma-} &= \boldsymbol{\Lambda}_{\sigma-} + \boldsymbol{\Lambda}_{\sigma 0}(-\boldsymbol{\Lambda}_{00})^{-1}\boldsymbol{\Lambda}_{0-}, \end{aligned}$$

$$\begin{aligned} \boldsymbol{\Lambda}^c_{+\sigma} &= \boldsymbol{\Lambda}_{+\sigma} + \boldsymbol{\Lambda}_{+0}(-\boldsymbol{\Lambda}_{00})^{-1}\boldsymbol{\Lambda}_{0\sigma}, \\ \boldsymbol{\Lambda}^c_{\sigma+} &= \boldsymbol{\Lambda}_{++} + \boldsymbol{\Lambda}_{+0}(-\boldsymbol{\Lambda}_{00})^{-1}\boldsymbol{\Lambda}_{0+}, \\ \boldsymbol{\Lambda}^c_{+-} &= \boldsymbol{\Lambda}_{+-} + \boldsymbol{\Lambda}_{+0}(-\boldsymbol{\Lambda}_{00})^{-1}\boldsymbol{\Lambda}_{0-}, \end{aligned}$$

$$\begin{aligned} \boldsymbol{\Lambda}^c_{-\sigma} &= \boldsymbol{\Lambda}_{-\sigma} + \boldsymbol{\Lambda}_{-0}(-\boldsymbol{\Lambda}_{00})^{-1}\boldsymbol{\Lambda}_{0\sigma}, \\ \boldsymbol{\Lambda}^c_{-+} &= \boldsymbol{\Lambda}_{-+} + \boldsymbol{\Lambda}_{-0}(-\boldsymbol{\Lambda}_{00})^{-1}\boldsymbol{\Lambda}_{0+}, \\ \boldsymbol{\Lambda}^c_{--} &= \boldsymbol{\Lambda}_{--} + \boldsymbol{\Lambda}_{-0}(-\boldsymbol{\Lambda}_{00})^{-1}\boldsymbol{\Lambda}_{0-}. \end{aligned}$$

Define

$$V_t^c = \int_0^t r_{J_s^c}\mathrm{d}s + \int_0^t \sigma_{J_s^c}\mathrm{d}B_s, \quad t \geq 0,$$

where $\{B_t\}_{t\geq 0}$ denotes an independent standard Brownian motion. See Figure 2.8 for an example of a $\mathfrak{E}_0$–censored process. Since $\{(V_t^c, J_t^c)\}_{t\geq 0}$ is such that $\mathfrak{E}_0 = \emptyset$, then its downcrossings and upcrossings probabilities can be studied through the formulae of Subsection 2.4.1 by replacing the intensity matrix $\boldsymbol{\Lambda}$ with $\boldsymbol{\Lambda}^c$. Furthermore, the first passage probabilities of $\{(V_t, J_t)\}_{t\geq 0}$ are exactly the same as the ones of $\{(V_t^c, J_t^c)\}_{t\geq 0}$, so that the formulae of Theorems 2.31, 2.32, 2.33, 2.34, 2.35 and 2.36 with $\boldsymbol{\Lambda}$ replaced by $\boldsymbol{\Lambda}^c$ also correspond to the downcrossing and upcrossing probabilities of the general fluid flow process $\{(V_t, J_t)\}_{t\geq 0}$.

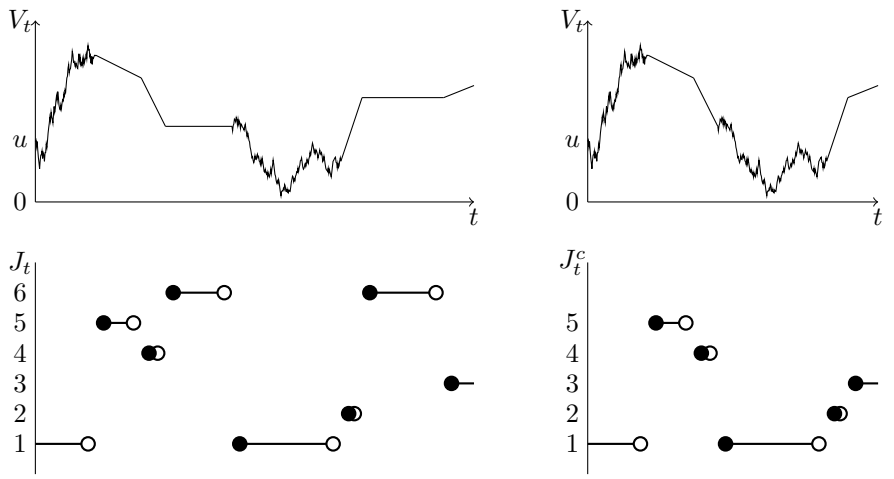


Figure 2.8: An example of a fluid flow process $\{(V_t, J_t)\}_{t\geq 0}$ (left) and its $\mathfrak{E}_0$–censored version $\{(V_t^c, J_t^c)\}_{t\geq 0}$ (right), where $\mathfrak{E}_\sigma = \{1\}$, $\mathfrak{E}_+ = \{2,3\}$, $\mathfrak{E}_- = \{4,5\}$ and $\mathfrak{E}_0 = \{6\}$. The holding times of $\{J_t\}_{t\geq 0}$ while at $\mathfrak{E}_0$ are eliminated in $\{J_t^c\}_{t\geq 0}$. Notice that $\{V_t\}_{t\geq 0}$ and $\{V_t^c\}_{t\geq 0}$ have the same first passage probabilities.

2.5 Risk processes

In this section we setup several continuous–time stochastic processes $\{R_t\}_{t\geq 0}$ which have been used widely used to model the capital of insurance companies, called risk processes. Rolski et al. (2009) and Asmussen and Albrecher (2010) constitute classic monographs on the study of these models. In the following, we focus our attention in five of these models: the Cramér–Lundberg process, the spectrally negative Lévy process, the Sparre–Andersen process, the risk process with MAP arrivals, and the fluid flow risk process.

2.5.1 Cramér–Lundberg process

The following process, called Cramér–Lundberg process, was intially formulated in Cramér (1930). Let $\{N_t\}_{t\geq 0}$ be a Poison process of rate $\lambda > 0$ and let $\{U_i\}_{i\geq 1}$ be an i.i.d. sequence of nonnegative random variables with common distribution F independent of $\{N_t\}_{t\geq 0}$. We define the **Cramér–Lundberg process** $\{R_t\}$

with **initial capital** $u > 0$ and **premium rate** $p > 0$ by

$$R_t = u + pt - \sum_{i=1}^{N_t} U_i,$$

where $\{N_t\}_{t\geq 0}$ represents the **claim arrival process** and $\{U_i\}_{i\geq 1}$ represents the sequence of **claim sizes**. See Figure 2.9 for an example of the path of a Cramér–Lundberg process $\{R_t\}_{t\geq 0}$. As in most risk–reserve process, we are

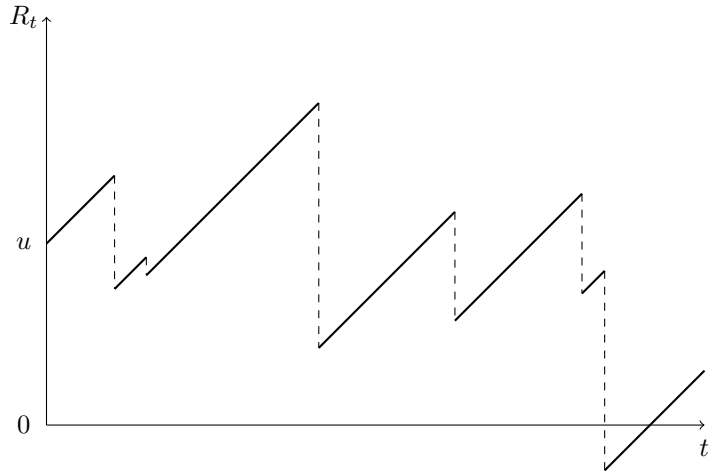


Figure 2.9: An example of a Cramér–Lundberg process $\{R_t\}_{t\geq 0}$. Between jumps, $\{R_t\}_{t\geq 0}$ increases at rate $p > 0$. Discontinuities happen according to a Poisson process $\{N_t\}_{t\geq 0}$, and jump sizes correspond to claim sizes $\{U_i\}_{i\geq 1}$.

interested in computing its infinite–horizon probability of ruin, that is,

$$\psi(u) := \mathbb{P}\left(\inf_{t\geq 0}\{R_t\} < 0 \mid R_0 = u\right).$$

The probability of ruin is non–trivial only in the case $R_t \to +\infty$ as $t \to \infty$, which happens if and only if $\lambda\mu < p$ where $\mu := \mathbb{E}(U_1)$; in the case $\lambda\mu \geq p$ ruin is certain and thus $\psi(u) = 1$ for all $u \geq 0$. Fortunately, for the non–trivial case we have an explicit formula for $\psi(u)$, discovered in Pollaczek (1930) and Khinchin (1967), which was formulated in a queueing setting related to the Cramér–Lundberg process.

Theorem 2.37 (Pollaczek–Khinchin formula) *Let $\{R_t\}_{t\geq 0}$ be a Cramér–Lundberg process as described above. Then,*

$$1 - \psi(u) = (1-\rho)\sum_{i=0}^{\infty} \rho^i F_{\mathrm{e}}^{*i}(u), \tag{2.5.1}$$

where $\rho = \lambda\mu/p$ *and* $F_{\mathrm{e}}(x) = \mu^{-1}\int_0^x (1 - F(s))\mathrm{d}s$.

Although explicit, formula (2.5.1) is not always easily computable. In most cases F_{e}^{*n} has a complex form, and even if it does not, we would need to compute an infinite sum involving F_{e}^{*n} for all $n \geq 1$. Notice that the r.h.s. of (2.5.1) has the following interpretation: if $\{Y_i\}_{i\geq 1}$ is an i.i.d. sequence of random variables with common distribution F_{e} and $Z \sim \mathrm{Geo}(1-\rho)$ with support on $\{0, 1, 2, \dots\}$, then

$$\mathbb{P}\left(\sum_{i=1}^{Z} Y_i \leq u\right) = (1-\rho)\sum_{i=0}^{\infty} \rho^i F_{\mathrm{e}}^{*i}(u). \tag{2.5.2}$$

Now, suppose that F corresponds to $\mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$. Then $\mu = \boldsymbol{\pi}(-\boldsymbol{T})^{-1}\boldsymbol{e}$ and

$$\begin{aligned} 1 - F_{\mathrm{e}}(x) &= \mu^{-1}\int_x^{\infty} (1 - F(s))\mathrm{d}s \\ &= (\boldsymbol{\pi}(-\boldsymbol{T})^{-1}\boldsymbol{e})^{-1}\int_x^{\infty} \boldsymbol{\pi} e^{\boldsymbol{T}s}\boldsymbol{e}\mathrm{d}s \\ &= (\boldsymbol{\pi}(-\boldsymbol{T})^{-1}\boldsymbol{e})^{-1}\boldsymbol{\pi}\left(\int_x^{\infty} e^{\boldsymbol{T}s}\mathrm{d}s\right)\boldsymbol{e} \\ &= \frac{\boldsymbol{\pi}(-\boldsymbol{T})^{-1}}{\boldsymbol{\pi}(-\boldsymbol{T})^{-1}\boldsymbol{e}} e^{\boldsymbol{T}x}\boldsymbol{e}, \end{aligned} \tag{2.5.3}$$

so that F_{e} corresponds to $\mathrm{PH}(\boldsymbol{\pi}_{\mathrm{e}}, \boldsymbol{T})$ with $\boldsymbol{\pi}_{\mathrm{e}} = \frac{\boldsymbol{\pi}(-\boldsymbol{T})^{-1}}{\boldsymbol{\pi}(-\boldsymbol{T})^{-1}\boldsymbol{e}}$. Equations (2.5.2) and (2.5.1), and Theorem 2.14 yield the following.

Theorem 2.38 *If* $\psi(u)$ *corresponds to the ruin probability of a Cramér–Lundberg process* $\{R_t\}_{t\geq 0}$ *with arrival rate* $\lambda > 0$*, premium rate* $p > 0$*,* $\mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$*–distributed claim sizes and initial capital* $u \geq 0$*, then*

$$\psi(u) = \begin{cases} \rho\boldsymbol{\pi}_{\mathrm{e}}\exp((\boldsymbol{T} + \rho\boldsymbol{t}\boldsymbol{\pi}_{\mathrm{e}})u)\boldsymbol{e} & \text{if} \quad \rho < 1 \\ 1 & \text{if} \quad \rho \geq 1, \end{cases}$$

where $\rho = \lambda\mu/p$*,* $\mu = \boldsymbol{\pi}(-\boldsymbol{T})^{-1}\boldsymbol{e}$ *and* $\boldsymbol{\pi}_{\mathrm{e}} = \frac{\boldsymbol{\pi}(-\boldsymbol{T})^{-1}}{\boldsymbol{\pi}(-\boldsymbol{T})^{-1}\boldsymbol{e}}$.

Thus, in the case the claims follow a phase–type distribution, the probability of ruin for the Cramér–Lundberg process will be easily computable. We will use this fact in Chapter 3 to propose a method to approximate the probability of ruin of Cramér–Lundberg processes with arbitrary claim size distribution.

2.5.2 Spectrally negative Lévy process

A Lévy process is usually thought as the continuous–time analogue of a random walk. Since its inception, the development of the theory of Lévy processes has been considerably satisfactory. Seminal books on the topic are Bertoin (1998) and Sato (1999). Furthermore, advances in fluctuation theory of this class of processes helped spread its use in stochastic modelling; see Kyprianou (2014). We give a precise definition of a Lévy process next.

A Markov process $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}, \{\mathbb{P}_x\}_{x\in\mathbb{R}})$ with $\mathcal{F}_t = \sigma(\{X_s\}_{s\leq t})$ is said to be a **Lévy process** if $\{X_t\}_{t\geq 0}$ has cádlág paths and stationary independent increments. The latter means that for all $t, s \geq 0$ and under the measure $\mathbb{P}_0$,

- $X_{t+s} - X_t \sim X_s$, and
- $X_{t+s} - X_t \perp \mathcal{F}_t$.

Furthermore, the family $\{\mathbb{P}_x\}_{x\in\mathbb{R}}$ is such that for each $x \in \mathbb{R}$, the law of $\{X_t\}_{t\geq 0}$ under $\mathbb{P}_x$ is the same as the law of $\{X_t + x\}_{t\geq 0}$ under $\mathbb{P}_0$. From now on, denote $\mathbb{P}_0$ by $\mathbb{P}$.

The following provides a characterisation of all Lévy processes, due to Lévy (1934) and Khintchine (1937).

Theorem 2.39 (Lévy–Khinchine formula) *Let $\{X_t\}_{t\geq 0}$ be a Lévy process and for all $\theta \in \mathbb{R}$ define the characteristic exponent $\Psi(\theta) := -\log \mathbb{E}(e^{\mathrm{i}\theta X_1})$. Then,*

$$\mathbb{E}(e^{\mathrm{i}\theta X_t}) = e^{-t\Psi(0)},$$

and $\Psi(\cdot)$ is on the form

$$\Psi(\theta) = \mathrm{i}a\theta + \frac{1}{2}\sigma^2\theta^2 + \int_{\mathbb{R}} (1 - e^{\mathrm{i}\theta x} + \mathrm{i}\theta x \mathbb{1}_{|x|<1})\Pi(\mathrm{d}x) \tag{2.5.4}$$

for some $a \in \mathbb{R}, \sigma \geq 0$ and Π a measure concentrated on $\mathbb{R}\setminus\{0\}$ such that $\int_{\mathbb{R}}(1 \wedge x^2)\Pi(\mathrm{d}x) < \infty$.

The following is the classic path decomposition of a Lévy process due to Lévy (1934) and Itô (1941).

Theorem 2.40 (Lévy–Itô decomposition) *If $\{X_t\}_{t\geq 0}$ is a Lévy process with characteristic exponent given by (2.5.4), then $\{X_t\}_{t\geq 0}$ can be decomposed as the sum of three independent processes $\{X_t^{(1)}\}_{t\geq 0}$, $\{X_t^{(2)}\}_{t\geq 0}$ and $\{X_t^{(3)}\}_{t\geq 0}$, where*

1. $\{X_t^{(1)}\}_{t\geq 0}$ *corresponds to a Brownian motion with drift which has characteristic exponent*
$$\Psi^{(1)}(\theta) = \mathrm{i}a\theta + \frac{1}{2}\sigma^2\theta^2,$$
2. $\{X_t^{(2)}\}_{t\geq 0}$ *corresponds to a compound Poisson process which has characteristic exponent*
$$\Psi^{(2)}(\theta) = \int_{|x|\geq 1} (1 - e^{\mathrm{i}\theta x})\Pi(\mathrm{d}x), \quad \textit{and}$$
3. $\{X_t^{(3)}\}_{t\geq 0}$ *corresponds to a square–integrable martingale with an almost surely countable number of jumps on each finite time interval, which has characteristic exponent*
$$\Psi^{(3)}(\theta) = \int_{0<|x|\leq 1} (1 - e^{\mathrm{i}\theta x} + \mathrm{i}\theta x)\Pi(\mathrm{d}x).$$

Throughout this manuscript we will focus on the family of Lévy processes with downward jumps only, denominated **spectrally negative Lévy processes**. See Figure 2.10 for an example of the path of a spectrally negative Lévy' process $\{R_t\}_{t\geq 0}$. For a spectrally negative Lévy process $\{X_t\}_{t\geq 0}$, the **Lévy expo-**

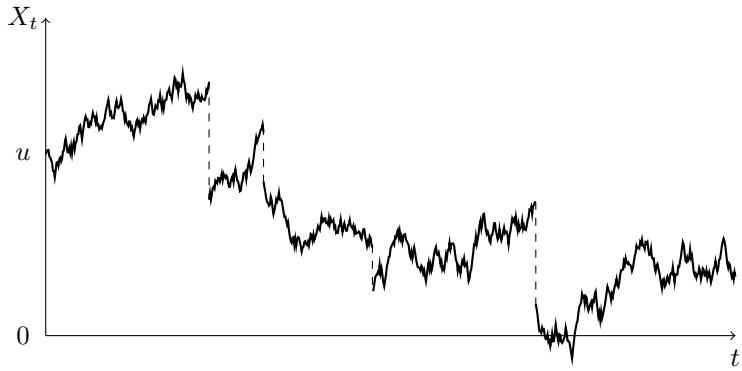


Figure 2.10: An example of a spectrally negative risk process $\{X_t\}_{t\geq 0}$ with $X_0 = 0$. This example has a Brownian component between jumps that happen at the arrival epochs of a Poisson process.

nent $\psi(\theta) := \log \mathbb{E}(e^{\theta X_1})$ is guaranteed to exist for all $\theta \geq 0$, and by analytic continuation of (2.5.4), it has the form

$$\psi(\theta) = a\theta + \frac{1}{2}\sigma^2\theta^2 + \int_{(-\infty,0)} (e^{\theta x} - 1 - \theta x \mathbb{1}_{x>-1})\Pi(\mathrm{d}x). \tag{2.5.5}$$

In the risk theory context, a spectrally negative Lévy process $\{X_t\}_{t\geq 0}$ such that $X_t \to \infty$ as $t \to \infty$ is sometimes called a **Lévy risk process**, following the reasoning that downward jumps model claim sizes, while the (possibly negative) premium is collected at a continuous rate. Notice that the condition $X_t \to \infty$ implies that $a > 0$ in (2.5.5). Furthermore, notice that the Cramér–Lundberg process is an example of a Lévy risk process.

2.5.3 Sparre–Andersen process

In Andersen (1957), a risk model is proposed as a generalization of the Cramér–Lundberg process, in the sense that the time between claims is no longer restricted to be exponential. The process, coined Sparre–Andersen process, will resemble a random walk, though, in general it does not belong to the class of Lévy process. We give its precise definition next.

Let $\{N_t\}_{t\geq 0}$ be a renewal process with interarrival times $\{S_i\}$ with common ditribution G and let $\{U_i\}_{i\geq 1}$ be an i.i.d. sequence of nonnegative random variables with common distribution F independent of $\{N_t\}_{t\geq 0}$. We define the **Sparre–Andersen process** $\{R_t\}$ with **initial capital** $u > 0$ and **premium rate** $p > 0$ by

$$R_t = u + pt - \sum_{i=1}^{N_t} U_i, \quad t \geq 0.$$

By taking $G \sim Exp(\lambda)$ $(\lambda > 0)$ one recovers the Cramér–Lundberg process. The general Sparre–Andersen process, however, lacks an explicit formula for its probability of ruin. A formula for the probability of ruin of a Sparre–Andersen process with phase–type and matrix–exponential components can be found in Asmussen and Rolski (1992) and Bladt et al. (1996), respectively.

2.5.4 Risk process with MAP arrivals

The following class of risk processes were inspired by Markov modulation. That is, models in which the "environment" (that is, intensity of arrivals or type of arrivals) changes according to an underlying Markov jump process; see Asmussen (1989) for details. In the following, we give a mathematical definition of how said modulation affects the risk process.

Let $\{(N_t, J_t)\}_{t\geq 0}$ be a $\text{MAP}_p(\boldsymbol{\alpha}, \boldsymbol{C}, \boldsymbol{D})$ where $\{J_t\}_{t\geq 0}$ has state–space $\mathfrak{E}$. Let $\{T_i\}_{i\geq 1}$ denote the arrival epochs of $\{N_t\}_{t\geq 0}$. Let $\{\bar{F}^{(k\ell)}\}\}_{k,\ell\in\{1,\ldots,p\}}$ be a col-

lection of nonnegative distribution functions. Let

$$R_t = u + \int_0^t r^*_{J_s} \mathrm{d}s + \int_0^t \sigma^*_{J_s} \mathrm{d}B_s + \sum_{i:T_i \le t} U_i^{(K_i L_i)}, \tag{2.5.6}$$

where $r^*_i, \sigma^*_i \ge 0$ for all $i \in \mathfrak{E}$, $\{B_t\}_{t\ge 0}$ is an independent standard Brownian motion, $K_i = J_{T_i^-}$, $L_i = J_{T_i}$, and for all $k, \ell \in \{1, \dots, p\}$, $\{U_i^{(k\ell)}\}_{i\ge 1}$ is an independent i.i.d. sequence with common distribution $F^{(k\ell)}$. That is, at each discontinuity of $\{N_t\}_{t\ge 0}$ induced by a jump from i to j in $\{J_t\}_{t\ge 0}$, we draw an independent claim size with distribution $F^{(ij)}$. We call $\{R_t\}_{t\ge 0}$ a **risk process with MAP arrivals**. See Figure 2.11 for an example of the path of $\{R_t\}_{t\ge 0}$. The class of risk processes with MAP arrivals corresponds to the class of spec-

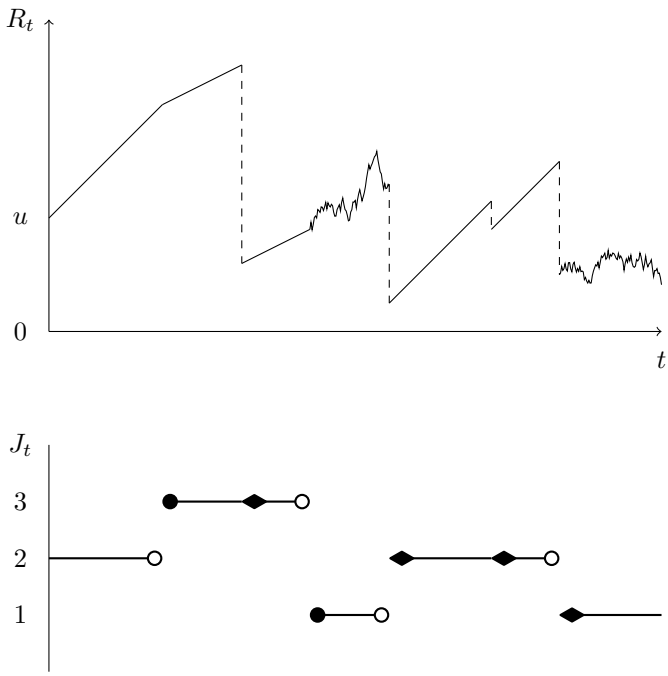


Figure 2.11: An example of a risk process $\{R_t\}_{t\ge 0}$ with MAP arrivals $\{(N_t, J_t)\}_{t\ge 0}$ with phase–space $\mathfrak{E}_\sigma \cup \mathfrak{E}_+ = \{1, 2, 3\}$. The arrival times are indicated in $\{J_t\}_{t\ge 0}$ with a diamond. In this example, the first claim was drawn according to the distribution $F^{(33)}$, the second one according to $F^{(12)}$, the third one according to $F^{(22)}$ and the fourth one according to $F^{(21)}$.

trally negative Markov additive processes with finite–jump activity; see Breuer (2008) and Ivanovs (2011).

Notice that if $G \sim \text{PH}(\boldsymbol{\pi}, \boldsymbol{T})$ and F is any general nonnegative distribution, then the risk process $\{R_t\}_{t\geq 0}$ with MAP arrivals $(\boldsymbol{\pi}, \boldsymbol{C}, \boldsymbol{D})$ with $\boldsymbol{C} = \boldsymbol{T}$, $\boldsymbol{D} = \boldsymbol{t}\boldsymbol{\pi}$ and $F^{(k\ell)} \sim F$ for all $k, \ell \in \{1, \ldots, p\}$ corresponds to a Sparre–Andersen process with $\text{PH}(\boldsymbol{\pi}, \boldsymbol{T})$–distributed interarrival times and F–distributed claim sizes.

2.5.5 Fluid flow risk process

This final class of risk processes is based on the following simple idea: for a given fluid flow process, regard its linear downward movements as downward jumps instead. This idea was first pursued in Asmussen (1995b). Since then, the use of fluid flow processes as risk processes has been advocated several times in the literature; see Badescu et al. (2005), Ren et al. (2009) and Breuer (2010) for some examples. The advantage of this class of risk processes, which we coin fluid flow risk processes in this thesis, is that several first pasage probabilities are directly inherited from the theory of fluid flow processes. Below we give the precise definition of a fluid flow risk process.

Let $\{(V_t, J_t)\}_{t\geq 0} = \{(V(t), J(t))\}_{t\geq 0}$ be a fluid flow process with phase–space $\mathfrak{E}_\sigma \cup \mathfrak{E}_+ \cup \mathfrak{E}_- \cup \mathfrak{E}_0$ such that $r_i = -1$ for all $i \in \mathfrak{E}_-$. Let $S_0 = U_0 = 0$ and for $i \geq 1$ define

$$S_i = \inf\left\{t > 0 : J\left(t + \sum_{k=1}^{i-1}(S_k + U_k)\right) \in \mathfrak{E}_-\right\},$$

$$U_i = \inf\left\{t > 0 : J\left(t + \sum_{k=1}^{i}(S_k + U_{k-1})\right) \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+ \cup \mathfrak{E}_0\right\}.$$

We define the **fluid flow risk process** $\{R_t\}_{t\geq 0}$ by

$$R_t = V\left(t + \sum_{k=1}^{i} U_k\right) \quad \text{for} \quad t \in \left[\sum_{k=1}^{i} S_k, \sum_{k=1}^{i+1} S_k\right),$$

and we call $\{(V_t, J_t)\}_{t\geq 0}$ its associated fluid flow process. See Figure 2.12 for an example of the path of a fluid flow risk process $\{R_t\}_{t\geq 0}$. Notice that $\{V_t\}_{t\geq 0}$ and $\{R_t\}_{t\geq 0}$ share the same first passage probabilities: this observation is the key behind computing the probability of ruin of $\{R_t\}_{t\geq 0}$. We will get back to this problem in Chapter 5.

Below we show how some risk models with phase–type–distributed claim sizes fit the framework of fluid flow risk processes.

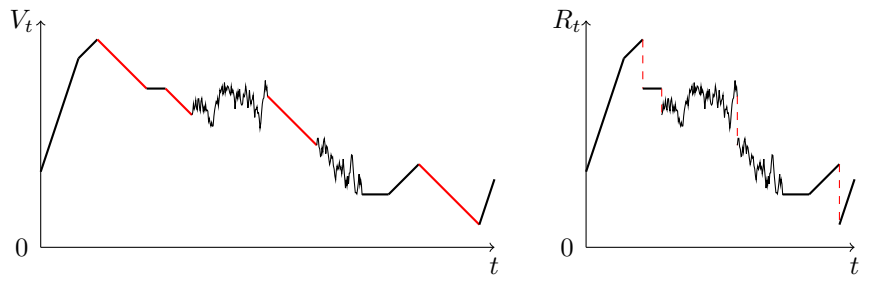


Figure 2.12: An example of a fluid flow process process $\{V_t\}_{t\geq 0}$ and its associated fluid flow risk process $\{R_t\}_{t\geq 0}$. Notice that the downward linear movements of $\{V_t\}_{t\geq 0}$ (shown in red) translate into jumps in $\{R_t\}_{t\geq 0}$.

Example 2.7 (Cramér–Lundberg processes) *Let $\{R_t\}_{t\geq 0}$ be the Cramér–Lundberg process with premium rate $p > 0$, initial capital, arrival rate $\lambda > 0$ and claim sizes with common distribution $\mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$. Let $\mathfrak{E}_T$ be the phase–space of $\mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$. Then $\{R_t\}_{t\geq 0}$ can be regarded as a fluid flow risk process with associated fluid flow process $\{(V_t, J_t)\}_{t\geq 0}$, where $\{J_t\}_{t\geq 0}$ has state–space $\mathfrak{E}_+ \cup \mathfrak{E}_-$ with $\mathfrak{E}_+ = \{1\}$ and $\mathfrak{E}_- = \mathfrak{E}_T$. Moreover, $\{J_t\}_{t\geq 0}$ has initial distribution $(1, \mathbf{0})$ and intensity matrix*

$$\begin{pmatrix} -\lambda & \lambda\boldsymbol{\pi} \\ \boldsymbol{t} & \boldsymbol{T} \end{pmatrix},$$

while the reward rates $\{r_i\}_{i\in\mathfrak{E}_+\cup\mathfrak{E}_-}$ and Brownian noise rates $\{r_i\}_{i\in\mathfrak{E}_+\cup\mathfrak{E}_-}$ are given by

$$\begin{aligned} r_i &= p \quad \text{and} \quad \sigma_i = 0, \quad i \in \mathfrak{E}_+, \\ r_i &= -1 \quad \text{and} \quad \sigma_i = 0, \quad i \in \mathfrak{E}_-. \end{aligned}$$

This embedding was originally proposed in Asmussen et al. (2002) in order to compute approximations to the finite–horizon probability of ruin of the Cramér–Lundberg process.

Example 2.8 (Spectrally negative Lévy processes) *Let $\{R_t\}$ be a spectrally negative Lévy process with $\mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$–distributed jumps. More precisely, $\{R_t\}_{t\geq 0}$ has Lévy exponent*

$$\psi(\theta) = p\theta + \frac{1}{2}\sigma^2\theta^2 + \int_{(-\infty,0)} (e^{\theta x} - 1 - \theta x \mathbb{1}_{x>-1})\Pi(\mathrm{d}x)$$

where $p > 0$ and $\Pi(\mathrm{d}x) = \lambda\boldsymbol{\pi}e^{\boldsymbol{T}x}\boldsymbol{t}$. The case $\sigma = 0$ corresponds to a Cramér–Lundberg process treated in Example 2.8. The case $\sigma > 0$ can be regarded as a fluid flow risk process with associated fluid flow process $\{(V_t, J_t)\}_{t\geq 0}$, where $\{J_t\}_{t\geq 0}$ has state–space $\mathfrak{E}_\sigma \cup \mathfrak{E}_-$ with $\mathfrak{E}_\sigma = \{1\}$ and $\mathfrak{E}_- = \mathfrak{E}_T$, where $\mathfrak{E}_T$

corresponds to the phase–state of $\mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$. *Furthermore,* $\{J_t\}_{t\geq 0}$ *has initial distribution* $(1, \mathbf{0})$ *and intensity matrix*

$$\begin{pmatrix} -\lambda & \lambda\boldsymbol{\pi} \\ \boldsymbol{t} & \boldsymbol{T} \end{pmatrix},$$

while the reward rates $\{r_i\}_{i\in\mathfrak{E}_\sigma\cup\mathfrak{E}_-}$ *and Brownian noise rates* $\{r_i\}_{i\in\mathfrak{E}_\sigma\cup\mathfrak{E}_-}$ *are given by*

$$\begin{aligned} r_i &= p \quad \textit{and} \quad \sigma_i = \sigma, \quad i \in \mathfrak{E}_\sigma, \\ r_i &= -1 \quad \textit{and} \quad \sigma_i = 0, \quad i \in \mathfrak{E}_-. \end{aligned}$$

This embedding was originally proposed in Asmussen et al. (2004) in order to compute some fluctuation identities of options based on phase–type Lévy models.

Example 2.9 (Sparre–Andersen processes) *Let* $\{R_t\}_{t\geq 0}$ *be a Sparre andersen process with premium rate* $p > 0$, $\mathrm{PH}(\boldsymbol{\alpha}, \boldsymbol{S})$*–distributed interarrival times and* $\mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$*–distributed claim sizes. Let* $\mathfrak{E}_S$ *and* $\mathfrak{E}_T$ *denote the phase–spaces of* $\mathrm{PH}(\boldsymbol{\alpha}, \boldsymbol{S})$ *and* $\mathrm{PH}(\boldsymbol{\pi}, \boldsymbol{T})$*, respectively. Then* $\{R_t\}$ *can be regarded as a fluid flow risk process with associated fluid flow process* $\{(V_t, J_t)\}_{t\geq 0}$*, where* $\{J_t\}_{t\geq 0}$ *has state–space* $\mathfrak{E}_+ \cup \mathfrak{E}_-$ *with* $\mathfrak{E}_+ = \mathfrak{E}_S$ *and* $\mathfrak{E}_- = \mathfrak{E}_T$*. The process* $\{J_t\}_{t\geq 0}$ *has an initial distribution* $(\boldsymbol{\alpha}, \mathbf{0})$ *and intensity matrix*

$$\begin{pmatrix} \boldsymbol{S} & \boldsymbol{s}\boldsymbol{\pi} \\ \boldsymbol{t}\boldsymbol{\alpha} & \boldsymbol{T} \end{pmatrix},$$

while the reward rates $\{r_i\}_{i\in\mathfrak{E}_+\cup\mathfrak{E}_-}$ *and Brownian noise rates* $\{r_i\}_{i\in\mathfrak{E}_+\cup\mathfrak{E}_-}$ *are given by*

$$\begin{aligned} r_i &= p \quad \textit{and} \quad \sigma_i = 0, \quad i \in \mathfrak{E}_+, \\ r_i &= -1 \quad \textit{and} \quad \sigma_i = 0, \quad i \in \mathfrak{E}_-. \end{aligned}$$

This embedding was first proposed in Stanford et al. (2005) in order to compute the finite–horizon probability of ruin of a Sparre–Andersen process.

Example 2.10 (Risk processes with MAP arrivals) *Let* $\{R_t\}$ *be a risk process with* $\mathrm{MAP}_p(\boldsymbol{\alpha}, \boldsymbol{C}, \boldsymbol{D})$*–arrivals defined as in (2.5.6). Moreover, suppose that for each* $i, j \in \{1, \ldots, p\}$, $F^{(ij)}$ *corresponds to a phase–type distribution with parameters* $(\boldsymbol{\pi}^{(ij)}, \boldsymbol{T}^{(ij)})$ *and phase–space* $\mathfrak{E}_{(ij)}$*. Then* $\{R_t\}_{t\geq 0}$ *can be regarded as a fluid flow risk process with underlying fluid flow process* $\{(V_t, J_t)\}_{t\geq 0}$*, where* $\{J_t\}_{t\geq 0}$ *has state–space* $(\mathfrak{E}_\sigma \cup \mathfrak{E}_+) \cup \mathfrak{E}_-$ *with*

$$\mathfrak{E}_\sigma \cup \mathfrak{E}_+ = \{1, \ldots, p\} \quad \textit{and} \quad \mathfrak{E}_- = \bigcup_{(i,j)\in\{1,\ldots,p\}^2} \mathfrak{E}_{(ij)}.$$

The process $\{J_t\}_{t\geq 0}$ *has initial distribution* $(\boldsymbol{\alpha}, \mathbf{0})$ *and intensity matrix*

$$\left(\begin{array}{ccc|ccccccc}
c_{11} & \cdots & c_{1p} & d_{11}\boldsymbol{\pi}^{(11)} & \cdots & d_{1p}\boldsymbol{\pi}^{(1p)} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\
c_{21} & \cdots & c_{2p} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
c_{p1} & \cdots & c_{pp} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & d_{11}\boldsymbol{\pi}^{(11)} & \cdots & d_{1p}\boldsymbol{\pi}^{(1p)} \\
\hline
\boldsymbol{t}^{(11)} & \cdots & \mathbf{0} & \boldsymbol{T}^{(11)} & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\mathbf{0} & \cdots & \boldsymbol{t}^{(1p)} & \mathbf{0} & \cdots & \boldsymbol{T}^{(1p)} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\boldsymbol{t}^{(p1)} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \boldsymbol{T}^{(p1)} & \cdots & \mathbf{0} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\mathbf{0} & \cdots & \boldsymbol{t}^{(pp)} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \boldsymbol{T}^{(pp)}
\end{array}\right),$$

while the reward rates $\{r_i\}_{i\in\mathfrak{E}_{\sigma+}\cup\mathfrak{E}_-}$ *and Brownian noise rates* $\{r_i\}_{i\in\mathfrak{E}_{\sigma+}\cup\mathfrak{E}_-}$ *are given by*

$$\begin{aligned} r_i &= r_i^* \quad \textit{and} \quad \sigma_i = \sigma_i^*, \quad i \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+, \\ r_i &= -1 \quad \textit{and} \quad \sigma_i = 0, \quad i \in \mathfrak{E}_-. \end{aligned}$$

This embedding was first explored in Ramaswami (2006) in order to compute some first passage probabilities of the risk process with MAP arrivals.

Chapter 3

An approximation of Cramér–Lundberg probability of ruin

3.1 Introduction

Consider a Cramér–Lundberg process $\{R_t\}_{t\geq 0}$ with arrival rate $\lambda > 0$, premium rate $p > 0$, F–distributed claim sizes and initial capital $u \geq 0$. From Theorem 2.37 we know that the probability of ruin $\psi(u)$ is equal to

$$1 - (1-\rho)\sum_{i=0}^{\infty} \rho^i F_{\mathrm{e}}^{*i}(u), \tag{3.1.1}$$

where $\rho = \lambda\mu/p$, μ is the mean of F and $F_{\mathrm{e}}(x) = \mu^{-1}\int_0^x (1-F(s))\mathrm{d}s$. Although exact, (3.1.1) is not easy to compute for general distributions F, except whenever F is phase–type–distributed (as noted in Theorem 2.38).

In this chapter we propose a method to approximate the probability of ruin of a Cramér–Lundberg process with arbitrary claim size distribution F which is either absolutely continuous or discrete, and has finite first moment. It is based on the approximation of an arbitrary distribution by a phase–type scale mixture;

see Vatamidou et al. (2012) and Bladt et al. (2014). We provide error bounds for our approximation and finalise by discusing the advantages and disavantages of our approach with respect to other methods in the literature.

The main original results stemming from this chapter are Theorem 3.1, Theorem 3.5 and Theorem 3.6.

3.2 Approximation of geometric random sums

For $u \geq 0$, $\rho \in (0,1)$ and distribution function H with support on $[0,\infty)$, define

$$\psi_{\mathrm{g}}(u,\rho,H) := 1-(1-\rho)\sum_{i=0}^{\infty}\rho^i H^{*i}(u). \tag{3.2.1}$$

According to (2.2.10), the r.h.s. of (3.2.1) can be interpreted as the distribution function of a geometric random sum of i.i.d.r.v.'s with common distribution H. Intituitively, if H_1 and H_2 are distribution functions which are "similar" to each other, then we expect $\psi_{\mathrm{g}}(u,\rho,H_1)$ to be close to $\psi_{\mathrm{g}}(u,\rho,H_2)$. The following is a precise statement of this fact, which is an improvement of the bound found in Proposition 2 of Vatamidou et al. (2012).

Theorem 3.1 *Let $u \geq 0$, $\rho \in (0,1)$, and let H_1 and H_2 be distribution functions with support on $[0,\infty)$. Then,*

$$|\psi_{\mathrm{g}}(u,\rho,H_1)-\psi_g(u,\rho,H_1)| \leq \sup_{s\leq u}\{|H_1(s)-H_2(s)|\}\,\frac{(1-\rho)\rho}{(1-\rho H_1(u))\,(1-\rho H_2(u))}. \tag{3.2.2}$$

Proof. First, let us verify that for all $n \geq 1$,

$$\sup_{s\leq u}\{|H_1^{*n}(s)-H_2^{*n}(s)|\} \leq \sup_{s\leq u}\{|H_1(s)-H_2(s)|\}\sum_{i=0}^{n-1}H_1^i(u)H_2^{n-1-i}(u). \tag{3.2.3}$$

We prove the previous by induction. It is clearly valid for $n=1$. Let us assume that it is valid for some $n \geq 1$. Then,

$$\begin{aligned}
&\sup_{s\leq u}\left\{\left|H_1^{*n+1}(s)-H_2^{*n+1}(s)\right|\right\}\\
&\quad=\sup_{s\leq u}\left\{\left|H_1^{*n+1}(s)-H_1^{*n}*H_2(s)+H_1^{*n}*H_2(s)-H_2^{*n+1}(s)\right|\right\}\\
&\quad\leq\sup_{s\leq u}\left\{\left|H_1^{*n+1}(s)-H_1^{*n}*H_2(s)\right|\right\}+\sup_{s\leq u}\left\{\left|H_1^{*n}*H_2(s)-H_2^{*n+1}(s)\right|\right\}.
\end{aligned} \tag{3.2.4}$$

Let us examine the summands in (3.2.4). We can bound the first one by

$$\begin{aligned}\sup_{s\le u}\left\{\left|H_1^{*n+1}(s)-H_1^{*n}*H_2(s)\right|\right\} &\le \sup_{s\le u}\left\{\int_0^s |H_1(r)-H_2(r)|\,\mathrm{d}H_1^{*n}(r)\right\}\\ &\le \sup_{s\le u}\left\{\int_0^s \sup_{\ell\le u}\{|H_1(\ell)-H_2(\ell)|\}\,\mathrm{d}H_1^{*n}(r)\right\}\\ &= \sup_{\ell\le u}\{|H_1(\ell)-H_2(\ell)|\}\sup_{s\le u}\left\{\int_0^s \mathrm{d}H_1^{*n}(r)\right\}\\ &= \sup_{\ell\le u}\{|H_1(\ell)-H_2(\ell)|\}\,H_1^{*n}(u)\\ &\le \sup_{\ell\le u}\{|H_1(\ell)-H_2(\ell)|\}\,H_1^{n}(u), \qquad (3.2.5)\end{aligned}$$

where in the last step we used the trivial inequality $H^{*n}\le H^n$ valid for all nonnegative distribution functions H. To bound the second summand of (3.2.4), we have that

$$\begin{aligned}&\sup_{s\le u}\left\{\left|H_1^{*n}*H_2(s)-H_2^{*n+1}(s)\right|\right\} \qquad (3.2.6)\\ &\quad\le \sup_{s\le u}\left\{\int_0^s |H_1^{*n}(r)-H_2^{*n}(r)|\,\mathrm{d}H_2(r)\right\}\\ &\quad\le \sup_{s\le u}\left\{\int_0^s \sup_{\ell\le u}\{|H_1^{*n}(\ell)-H_2^{*n}(\ell)|\}\,\mathrm{d}H_2(r)\right\}\\ &\quad= \sup_{\ell\le u}\{|H_1^{*n}(\ell)-H_2^{*n}(\ell)|\}\sup_{s\le u}\left\{\int_0^s \mathrm{d}H_2(r)\right\}\\ &\quad\le \left(\sup_{\ell\le u}\{|H_1(\ell)-H_2(\ell)|\}\sum_{i=0}^{n-1}H_1^i(u)H_2^{n-1-i}(u)\right)H_2(u)\\ &\quad= \sup_{\ell\le u}\{|H_1(\ell)-H_2(\ell)|\}\sum_{i=0}^{n-1}H_1^i(u)H_2^{n-i}(u), \qquad (3.2.7)\end{aligned}$$

where in the second–to–last inequality we used the induction hypothesis. Summing (3.2.5) and (3.2.7) we get that

$$\sup_{s\le u}\left\{\left|H_1^{*n+1}(s)-H_2^{*n+1}(s)\right|\right\}\le \sup_{s\le u}\{|H_1(s)-H_2(s)|\}\sum_{i=0}^{n}H_1^i(u)H_2^{n-i}(u),$$

so that (3.2.3) is valid for all $n \geq 1$. Finally,

$$
\begin{aligned}
&|\psi_{\mathrm{g}}(u,\rho,H_1) - \psi_{\mathrm{g}}(u,\rho,H_2)| \\
&\leq \sum_{n=1}^{\infty}(1-\rho)\rho^n \left|{H_1}^{*n}(u) - {H_2}^{*n}(u)\right| \\
&\leq \sup_{s\leq u}\{|H_1(s)-H_2(s)|\}\,(1-\rho)\sum_{n=1}^{\infty}\rho^n\sum_{i=0}^{n-1}{H_1}^{i}(u){H_2}^{n-1-i}(u) \\
&= \sup_{s\leq u}\{|H_1(s)-H_2(s)|\}\,(1-\rho)\sum_{i=0}^{\infty}\sum_{n=i+1}^{\infty}\rho^n{H_1}^{i}(u){H_2}^{n-1-i}(u) \\
&= \sup_{s\leq u}\{|H_1(s)-H_2(s)|\}\,(1-\rho)\sum_{i=0}^{\infty}\sum_{n=0}^{\infty}\rho^{n+i+1}{H_1}^{i}(u){H_2}^{n}(u) \\
&= \sup_{s\leq u}\{|H_1(s)-H_2(s)|\}\,(1-\rho)\rho\sum_{i=0}^{\infty}\rho^i{H_1}^{i}(u)\sum_{n=0}^{\infty}\rho^n{H_2}^{n}(u) \\
&= \sup_{s\leq u}\{|H_1(s)-H_2(s)|\}\,\frac{(1-\rho)\rho}{(1-\rho H_1(u))\,(1-\rho H_2(u))}.
\end{aligned}
$$

□

Our bound improves the one in Proposition 2 of Vatamidou et al. (2012) by a factor of

$$\frac{(1-\rho)^2}{(1-\rho H_1(u))\,(1-\rho H_2(u))}. \tag{3.2.8}$$

However, notice that as $u \to \infty$, (3.2.8) goes to 1, so that for large values of u, the bound of Proposition 2 of Vatamidou et al. (2012) and the one provided by Theorem 3.1 are close to each other.

3.3 Two kinds of approximations

For a given nonnegative distribution function F with finite mean μ, define the nonnegative distribution functions F_{e} and F_{m} given by

$$F_{\mathrm{e}}(x) = \mu^{-1}\int_0^x (1-F(s))\mathrm{d}s, \quad \text{and} \tag{3.3.1}$$

$$F_{\mathrm{m}}(x) = \mu^{-1}\int_0^x sF(\mathrm{d}s). \tag{3.3.2}$$

In the literature, F_{e} is called the **integrated tail distribution of** F, and F_{m} is called the **moment distribution of** F. While it is clear from (3.1.1) that F_{e}

plays an important role in the computation of $\psi(u)$, F_{m} also does so in a slightly subtler way. To see this, for nonnegative distributions F and G with finite first moments μ and ν, define their **Mellin convolution**, denoted by $H \star G$, by taking

$$(F \star G)(u) = \int_0^\infty G(u/s)F(\mathrm{d}s), \quad x \geq 0.$$

The Mellin convolution $F \star G$ has an easy intepretation: if $X \sim F$, $Y \sim G$ and X is independent of Y, then $XY \sim F \star G$. The previous probabilistic interpretation of the Mellin convolution implies that $F \star G = G \star F$, and also that the first moment of $F \star G$ is equal to $\mu\nu$. The following exhibits a link between F_{e} and F_{m} by means of the Mellin convolution.

Proposition 3.2 *Let F and G be nonnegative distributions with finite first moments. Then,*

$$(F \star G)_{\mathrm{e}} = F_{\mathrm{m}} \star G_{\mathrm{e}}. \tag{3.3.3}$$

In particular,

$$F_{\mathrm{e}} = F_{\mathrm{m}} \star U, \tag{3.3.4}$$

where U denotes a $U(0,1)$–distribution.

Proof. Let μ and ν denote the first moments of F and G, respectively. Observe that

$$\begin{aligned}
(F \star G)_{\mathrm{e}} &= \frac{1}{\mu\nu}\int_0^x (1 - (F \star G)(s))\,\mathrm{d}s \\
&= \frac{1}{\mu}\int_0^x \int_0^\infty \frac{1 - G(s/r)}{\nu} F(\mathrm{d}r)\mathrm{d}s \\
&= \int_0^\infty G_{\mathrm{e}}(x/r)\frac{r\mathrm{d}F(\mathrm{d}r)}{\mu} \\
&= \int_0^\infty G_{\mathrm{e}}(x/r)F_{\mathrm{m}}(\mathrm{d}r) \\
&= (G_{\mathrm{e}} \star F_{\mathrm{m}})(x) = (F_{\mathrm{m}} \star G_{\mathrm{e}})(x),
\end{aligned} \tag{3.3.5}$$

so that (3.3.3) follows. To prove (3.3.4), simply note that $F = F \star \delta_1$ and that

$$\begin{aligned}
(\delta_1)_{\mathrm{e}}(x) &= \int_0^x (1 - \delta_1(s))\mathrm{d}s \\
&= \int_0^x (1 - \mathbb{1}_{[1,\infty)}(s))\mathrm{d}s \\
&= \int_0^x (\mathbb{1}_{[0,1)}(s))\mathrm{d}s,
\end{aligned}$$

from which we can conclude that $(\delta_1)_{\mathrm{e}}$ corresponds to the $U(0,1)$ distribution and the proof is finished. In the following, we discuss two ways of approximating the probability of ruin for the Cramér–Lundberg process based on Theorem 3.1 and approximations of F_{e} and F_{m}.

3.3.1 Approximating F_{e}

Suppose we have a closed–form formula for F and its mean μ. This means that the density function associated to F_{e} is also closed–form, given by $f_{\mathrm{e}} = \mu^{-1}(1-F)$. With this in mind, we can approximate the probability of ruin

$$\psi(u) = 1 - (1-\rho)\sum_{i=0}^{\infty} \rho^i F_{\mathrm{e}}^{*i}(u)$$

as follows.

1. We can fit a phase–type density h to f_{e} following the algorithm of Asmussen et al. (1996). Once h is determined to correspond to the distribution $H \sim \mathrm{PH}(\boldsymbol{\alpha}, \boldsymbol{S})$, we can use the approximation

$$\psi(u,\rho,H) = 1 - (1-\rho)\sum_{i=0}^{\infty} \rho^i H^{*i}(u),$$

   which by (2.2.10) we know can be expresed as

$$\psi(u,\rho,H) = \rho\boldsymbol{\alpha} e^{(\boldsymbol{S}+\rho\boldsymbol{s}\boldsymbol{\alpha})u}\boldsymbol{e}.$$

2. We can fit a probability mass function h to f_{e} with support on $\{0, r, 2r, 3r, \dots\}$, for some fixed $r > 0$. Once the masses $\{h(ir)\}_{i=0}^{\infty}$ and the associated distribution function H are determined, we can follow the recursive algorithm of Panjer (1986) to compute

$$\psi(u,\rho,H) = 1 - (1-\rho)\sum_{i=0}^{\infty} \rho^i H^{*i}(u),$$

   which is given by the formula

$$\psi(u,\rho,H) = \psi_{\mathrm{app}}(\lfloor u/r \rfloor r, \rho, H)$$

with

$$\psi_{\text{app}}(ir,\rho,H) = \frac{1-H(ir)}{(1-\rho)^{-1}-h(0)} - \frac{1}{(1-\rho)^{-1}-h(0)}\sum_{j=1}^{i} h(jr)\psi_{\text{app}}((i-j)r,\rho,H),$$

$$\psi_{\text{app}}(0,\rho,H) = \frac{1-h(0)}{(1-\rho)^{-1}-h(0)}.$$

In either case, the approximation $\psi(u,\rho,H)$ will be such that

$$|\psi(u)-\psi(u,\rho,H)| = |\psi(u,\rho,F_{\text{e}})-\psi(u,\rho,H)| \le \sup_{s\le u}\{|F_{\text{e}}(s)-H(s)|\}\,\frac{(1-\rho)\rho}{(1-\rho F_{\text{e}}(u))\,(1-\rho H(u))}, \quad (3.3.6)$$

according to Theorem 3.1.

3.3.2 Approximating F_{m}

In this subsection, suppose that there is no closed–form formula for F, F is either absolutely continuous or discrete, and its density or mass function f has a closed–form formula. By (3.3.2), F_{m} has a density or mass function f_{m} given by

$$f_{\text{m}}(x) = \mu^{-1}xf(x),$$

where x varies over the support of f. Notice that if F is absolutely continuous (discrete), then F_m is absolutely continuous (discrete).

If we were to approximate F_{m} with, say, a distribution function H, how well does $H \star U$ approximate F_{e}? According to Proposition 3.2, they should not be too different; we give a precise statement next.

Lemma 3.3 *For any nonnegative distribution functions F and H,*

$$\sup_{s\le u}\{|F_{\text{e}}(s)-(H\star U)(s)|\} \le \sup_{s}\{|F_{\text{m}}(s)-H(s)|\}.$$

PROOF. Simply notice that

$$\begin{aligned}\sup_{s\le u}\{|F_{\mathrm{e}}(s)-(H\star U)(s)|\} &= \sup_{s\le u}\{|(F_{\mathrm{m}}\star U)(s)-(H\star U)(s)|\}\\ &= \sup_{s\le u}\left\{\left|\int_0^1 F_{\mathrm{m}}(s/r)-H(s/r)\mathrm{d}r\right|\right\}\\ &\le \sup_{s\le u}\left\{\int_0^1 |F_{\mathrm{m}}(s/r)-H(s/r)|\,\mathrm{d}r\right\}\\ &\le \int_0^1 \sup_s\{|F_{\mathrm{m}}(s/r)-H(s/r)|\}\,\mathrm{d}r\\ &= \sup_s\{|F_{\mathrm{m}}(s)-H(s)|\}.\end{aligned}$$

Lemma 3.3 and Theorem 3.1 imply that

$$|\psi(u)-\psi(u,\rho,H\star U)|\le \sup_s\{|F_{\mathrm{m}}(s)-H(s)|\}\,\frac{(1-\rho)\rho}{(1-\rho F_{\mathrm{e}}(u))\,(1-\rho(H\star U)(u))}. \tag{3.3.7}$$

Thus, $\psi(u,\rho,H\star U)$ constitutes an approximation to the probability of ruin whenever H and F_{m} are close enough. The downside is that there is no general direct method to compute $\psi(u,\rho,H\star U)$. We propose a method to **approximate** $\psi(u,\rho,H\star U)$ whenever H corresponds to a positive discrete distribution with a finite amount of atoms, say at $\{r_i\}_{i=1}^N$. Such method is based on the following observation.

Proposition 3.4 *Let E^k denote the Erlang distribution of mean 1 and dimension k, that is, the distribution with density function g given by*

$$g(x)=ke^{-kx}\frac{(kx)^{k-1}}{(k-1)!},\quad x\ge 0.$$

Then,

$$E^k \xrightarrow{w} \delta_1 \quad as \quad k\to\infty, \qquad and \tag{3.3.8}$$

$$E_{\mathrm{e}}^k \xrightarrow{w} U \quad as \quad k\to\infty, \tag{3.3.9}$$

where $\xrightarrow{w}$ denotes weak convergence. Moreover,

$$1-E_{\mathrm{e}}^k(1)=\mathrm{e}^{-k}\frac{k^{k-1}}{(k-1)!}\le (2\pi k)^{-\frac{1}{2}}. \tag{3.3.10}$$

PROOF. Let $X_k\sim E^k$. It is known that $\mathrm{Var}(X_k)=\mathbb{E}((X_k-1)^2)=1/k$, so that X_k converges in L_2 to 1, which in turn implies convergence in distribution

and (3.3.8) follows. Equation (3.3.9) follows by noting that (3.3.8) implies the a.e. pointwise convergence of the density function of E_{e}^k to the density function of U.

To prove (3.3.10), notice that

$$\begin{aligned}
1 - E_{\mathrm{e}}^k(1) &= 1 - \int_0^1 (1 - E^k(s))\mathrm{d}s = \int_0^1 E^k(s)\mathrm{d}s \\
&= \int_0^1 \left(1 - \sum_{n=0}^{k-1} \frac{1}{n!}\mathrm{e}^{-ks}(ks)^n\right)\mathrm{d}s = 1 - \sum_{n=0}^{k-1} \frac{1}{n!}\int_0^1 \mathrm{e}^{-ks}(ks)^n\mathrm{d}s \\
&= 1 - \sum_{n=0}^{k-1} \frac{1}{n!}\left(n!k^{-1} - \mathrm{e}^{-k}\sum_{k=0}^{n}\frac{n!k^{k-1}}{k!}\right) = \mathrm{e}^{-k}\sum_{n=0}^{k-1}\sum_{k=0}^{n}\frac{k^{k-1}}{k!} \\
&= \mathrm{e}^{-k}\sum_{k=0}^{k-1}(k-k)\frac{k^{k-1}}{k!} = \mathrm{e}^{-k}\left(\sum_{k=0}^{k-1}\frac{k^k}{k!} - \sum_{k=0}^{k-1}k\frac{k^{k-1}}{k!}\right) \\
&= \mathrm{e}^{-k}\left(\sum_{k=0}^{k-1}\frac{k^k}{k!} - \sum_{k=1}^{k-1}\frac{k^{k-1}}{(k-1)!}\right) \\
&= \mathrm{e}^{-k}\left(\sum_{k=0}^{k-1}\frac{k^k}{k!} - \sum_{k=0}^{k-2}\frac{k^k}{k!}\right) = \mathrm{e}^{-k}\frac{k^{k-1}}{(k-1)!}.
\end{aligned}$$

Finally, an application of Stirling's formula $k! > \sqrt{2\pi}k^{k+\frac{1}{2}}\mathrm{e}^{-k}$ yields $1-\widehat{E}^k(1) < (2\pi k)^{-\frac{1}{2}}$ and the proof is finished. □

Remark 11 *Not only does E_k converge weakly to δ_1 as k goes to infinity: amongst the class of phase–type distributions of dimension k, E_k is the one that approximates δ_1 the best. This statement was proved in Aldous and Shepp (1987) and in O'Cinneide (1991b); in the former it is checked that E_k is the least variable distribution amongst* PH_k *in a coefficient of variation sense, and in the latter it was checked that E_k is the least variable distribution amongst* PH_k *in a convex order sense. This is what inspired the concept of* **erlangization***: approximating a deterministic variable with a weakly convergent sequence of Erlang distributed random variables.*

Proposition 3.4 implies that for sufficiently large $k \geq 0$, $\psi(u, \rho, H \star E_{\mathrm{e}}^k)$ must be a "good" approximation of $\psi(u, \rho, H \star U)$. We assess the quality of such an approximation next.

Theorem 3.5

$$\left|\psi(u,\rho,H\star U)-\psi(u,\rho,H\star E_{\mathrm{e}}^{k})\right|\leq\frac{\rho\left(1-E_{\mathrm{e}}^{k}(1)\right)}{1-\rho E_{\mathrm{e}}^{k}(1)}\leq\frac{\rho}{1-\rho}\sqrt{\frac{1}{2\pi k}}.$$

Proof. Let $\{X_n\}_{n\geq 1}$ be a sequence i.i.d.r.v.'s with common distribution H. Then,

$$\begin{aligned}
&\left|(H\star U)^{*n}(u)-(H\star E_{\mathrm{e}}^{k})^{*n}(u)\right|\\
&\quad=\left|\int_{\mathbb{R}_+^n}\cdots\int\mathbb{P}(s_1X_1+\cdots+s_nX_n\leq u)U(\mathrm{d}s_1)\ldots U(\mathrm{d}s_n)\right.\\
&\qquad\left.-\int_{\mathbb{R}_+^n}\cdots\int\mathbb{P}(s_1X_1+\cdots+s_nX_n\leq u)E_{\mathrm{e}}^{k}(\mathrm{d}s_1)\ldots E_{\mathrm{e}}^{k}(\mathrm{d}s_n)\right|\\
&\quad\leq\left|\int_{\mathbb{R}_+^n}\cdots\int\begin{array}{c}\mathbb{P}(s_1X_1+\cdots+s_nX_n\leq u)\\ \times\left(\prod_{i=1}^n\mathbb{1}_{[0,1)}(s_i)-\prod_{i=1}^n(1-E^k(s_i))\right)\end{array}\mathrm{d}s_1\ldots\mathrm{d}s_n\right|\\
&\quad=\left|\int_{[0,1]^n}\cdots\int\begin{array}{c}\mathbb{P}(s_1X_1+\cdots+s_nX_n\leq u)\\ \times\left(\prod_{i=1}^n\mathbb{1}_{[0,1)}(s_i)-\prod_{i=1}^n(1-E^k(s_i))\right)\end{array}\mathrm{d}s_1\ldots\mathrm{d}s_n\right.\\
&\qquad\left.-\int_{\mathbb{R}^n\setminus[0,1]^n}\cdots\int\begin{array}{c}\mathbb{P}(s_1X_1'+\cdots+s_nX_n\leq u)\\ \times\left(\prod_{i=1}^n(1-E^k(s_i))\right)\end{array}\mathrm{d}s_1\ldots\mathrm{d}s_n\right|\\
&\quad\leq\sup\left\{\begin{array}{c}\int_{[0,1]^n}\cdots\int\left(\prod_{i=1}^n\mathbb{1}_{[0,1)}(s_i)-\prod_{i=1}^n(1-E^k(s_i))\right)\mathrm{d}s_1\ldots\mathrm{d}s_n\ ,\\ \int_{\mathbb{R}_+^n\setminus[0,1]^n}\cdots\int\prod_{i=1}^n(1-E^k(s_i))\mathrm{d}s_1\ldots\mathrm{d}s_n\end{array}\right\}.
\end{aligned}$$

Notice that

$$\begin{aligned}
&\int_{[0,1]^n}\cdots\int\left(\prod_{i=1}^n\mathbb{1}_{[0,1)}(s_i)-\prod_{i=1}^n(1-E^k(s_i))\right)\mathrm{d}s_1\ldots\mathrm{d}s_n\\
&\qquad=1-\left(\int_0^1(1-E^k(s)\mathrm{d}s\right)^n=1-(E_{\mathrm{e}}^{k}(1))^n,
\end{aligned}$$

and that

$$\int_{\mathbb{R}_+^n\setminus[0,1]^n}\cdots\int\prod_{i=1}^n(1-E^k(s_i))\mathrm{d}s_1\ldots\mathrm{d}s_n$$
$$=1-\int_{[0,1]^n}\cdots\int\prod_{i=1}^n(1-E^k(s_i))\mathrm{d}s_1\ldots\mathrm{d}s_n=1-(E_{\mathrm{e}}^k(1))^n,$$

so that

$$\left|(H\star U)^{*n}(u)-(H\star E_{\mathrm{e}}^k)^{*n}(u)\right|\leq 1-(E_{\mathrm{e}}^k(1))^n.$$

Therefore, we have that

$$\begin{aligned}
&\left|\psi(u,\rho,H\star U)-\psi(u,\rho,H\star E_{\mathrm{e}}^k)\right|\\
&\quad\leq\sum_{n=1}^{\infty}(1-\rho)\rho^n\left|(H\star U)^{*n}(u)-(H\star E_{\mathrm{e}}^k)^{*n}(u)\right|\\
&\quad\leq\sum_{n=1}^{\infty}(1-\rho)\rho^n\left(1-(E_{\mathrm{e}}^k(1))^n\right)\\
&\quad=\rho\left(1-\frac{(1-\rho)E_{\mathrm{e}}^k(1)}{1-\rho E_{\mathrm{e}}^k(1)}\right)\\
&\quad=\frac{\rho\left(1-E_{\mathrm{e}}^k(1)\right)}{1-\rho E_{\mathrm{e}}^k(1)}\leq\frac{\rho}{1-\rho}\sqrt{\frac{1}{2\pi k}},
\end{aligned}$$

and the proof is finished. □

The final step is to make sense of $H\star E_{\mathrm{e}}^k$. Since E^k follows an Erlang distribution of mean 1 and dimension k, its phase–type representation is given by $\mathrm{PH}_k(\boldsymbol{e}_1',\boldsymbol{E}_k)$ where

$$\boldsymbol{E}_k=\begin{pmatrix}-k & k & 0 & \cdots & 0\\ 0 & -k & k & \cdots & 0\\ 0 & 0 & -k & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & -k\end{pmatrix}.$$

Following (2.5.3), it can be verified that $E_{\mathrm{e}}^k\sim\mathrm{PH}_k(k^{-1}\boldsymbol{e}',\boldsymbol{E}_k)$. Thus, if H is a positive discrete distribution with atoms in $\{r_i\}_{i=1}^N$ of size $\{h_i\}_{i=1}^N$, then

$$H\star E_{\mathrm{e}}^k(x)=\sum_{i=1}^N h_iE_{\mathrm{e}}^k(x/r_i)$$

corresponds to a finite mixture of phase–type distributions. By (S2.2.3), $H \star E_{\mathrm{e}}^k$ is itself a phase–type distribution of parameters $(\boldsymbol{\alpha}_k^*, \boldsymbol{T}_k^*)$ with

$$\boldsymbol{\alpha}_k^* = k^{-1}(h_1\boldsymbol{e}', h_2\boldsymbol{e}', \ldots, h_N\boldsymbol{e}'), \quad \text{and} \tag{3.3.11}$$

$$\boldsymbol{T}_k^* = \begin{pmatrix} \boldsymbol{E}_k/r_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{E}_k/r_2 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{E}_k/r_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{E}_k/r_N \end{pmatrix}. \tag{3.3.12}$$

This means that $\psi(u, \rho, H \star E_{\mathrm{e}}^k)$ can be explicitly computed by means of (2.2.10). Thus, taking into account Theorem 2.38, (3.3.7) and Theorem 3.5 we get the following.

Theorem 3.6 *Let H be a postive discrete distribution with atoms in $\{r_i\}_{i=1}^N$ of size $\{h_i\}_{i=1}^N$. Then,*

$$|\psi(u) - \psi(u, \rho, H \star E_{\mathrm{e}}^k)| \leq \epsilon_1(F_{\mathrm{m}} - H) + \epsilon_2(k),$$

where

$$\psi(u, \rho, H \star E_{\mathrm{e}}^k) = \rho\boldsymbol{\alpha}_k^* \exp((\boldsymbol{T}_k^* + \rho\boldsymbol{t}_k^*\boldsymbol{\alpha}_k^*)u)\boldsymbol{e} \tag{3.3.13}$$

with $\boldsymbol{\alpha}^*$ and $\boldsymbol{T}_k^*$ given by (3.3.11) and (3.3.12), $\boldsymbol{t}_k^* := -\boldsymbol{T}_k^*\boldsymbol{e}$,

$$\epsilon_1(F_{\mathrm{m}} - H) = \sup_s \{|F_{\mathrm{m}}(s) - H(s)|\} \frac{(1-\rho)\rho}{(1-\rho F_{\mathrm{e}}(u))\,(1-\rho(H \star U)(u))}, \quad \textit{and}$$

$$\epsilon_2(k) = \frac{\rho}{1-\rho}\sqrt{\frac{1}{2\pi k}}.$$

Thus, Theorem 3.6 provides a way to approximate $\psi(u)$ with the closed–form formula given by (3.3.13). Its error bound is on the form $\epsilon_1 + \epsilon_2$, where ϵ_1 depends solely on the approximation error of H to F_{m}, while ϵ_2 only depends on the dimension used for the erlangization.

3.4 Conclusions and remarks

We provided a method to approximate the probability of ruin of a Crámer–Lundberg process. Our method compares to the existing methods in the literature as follows:

- Our methodology is based in calibration rather than statistical fitting. Since it is based on a brute–force approach of approximating distributions with a mixture of Erlang distributions, the dimension of the matrices required to make the computations is exceedingly large. If F_{e} is of closed–form, we believe that fitting F_{e} with a phase–type distribution as in Asmussen et al. (1996) is the most elegant and efficient way to go and the error bound of such an approximation will be given by our bond in (3.3.6).
- The approximation of Panjer (1986) based on discretization of F_{e} is straightforward to implement when F_{e} is of closed–form. The error bound of such an approximation will also be given by our bond in (3.3.6).
- The density or mass function of F_{m} is always of closed–form whenever the density or mass function of F is of closed–form. This is not always the case with F_{e}. Thus, Theorem 3.6 provides a general systematic way to approximate the probability of ruin of any Cramér–Lundberg process with discrete or continuous claim sizes, irrespective of F_{e} having a closed–form formula or not. Moreover, we give a bound $\epsilon_1 + \epsilon_2$ for the error of such an approximation.
- If F is discrete, then F_{m} is discrete. This means that in the case of discrete claim sizes with compact support, we can choose $H = F_{\mathrm{m}}$ and get $\epsilon_1 = 0$. If F_{m} has an infinite number of atoms, we would still able to choose $H = F_{\mathrm{m}}$ and compute the probability of ruin using the theory of infinite–dimensional phase–type distributions; this class of distributions may be used to approximate heavy tailed ones, see Bladt et al. (2014) and Peralta et al. (2018) for details. Otherwise, if we want to keep H within the class of finite–dimensional phase–type distributions we can do it, however, we would have that $\epsilon_1 > 0$.

Chapter 4

Bivariate phase–type distributions and their application to fluid flow risk processes

4.1 Introduction

In section 2.2.2 we learnt how to compute the distribution of the marginals of a given MPH* Kulkarni's multivariate phase–type distribution. In this chapter we focus on a class of bivariate distributions which are a subset of the class MPH* and are constructed according to the following principle: the bivariate distribution must have given marginals and given Pearson's correlation coefficient. In order to do so we follow the work of Baker (2008), where multivariate distributions with given marginals are constructed by using the order statistics of said marginals. Later on, we show how to use this class of bivariate distribution in order to model different kinds of dependent Sparre–Andersen processes with phase–type–distributed interarrival times and claim sizes. We finalise by discusing the benefits and disadvantages of our contruction with respect to other approaches in the literature.

The main original results stemming from this chapter are located in Section 4.3, Subsection 4.4.1, Subsection 4.4.2 and Subsection 4.4.3.

4.2 Bivariate distribution of the Baker–type

Let F denote any distribution function over $\mathbb{R}$ and for fixed $n \geq 1$, let $\{X_i\}_{i=1}^n$ be a collection of i.i.d.r.v.'s with common distribution F. Let $X_{i:n}$ denote the i-th order statistic among n, so that $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$. For each $i \in \{1, \ldots, n\}$, let $F_{i:n}$ denote the distribution function of $X_{i:n}$, so that

$$F_{i:n}(x) = \sum_{k=i}^{n} \binom{n}{k} (F(x))^k (1 - F(x))^{n-k}, \quad x \in \mathbb{R}.$$

Theorem 4.1 *Let F be any distribution function over $\mathbb{R}$ and let $\{F_{i:n}\}_{i=1}^n$ be their corresponding order statistics. Then,*

$$F = \frac{1}{n} \sum_{i=1}^{n} F_{i:n}. \tag{4.2.1}$$

Proof. For any $x \in \mathbb{R}$,

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^{n} F_{i:n}(x) &= \frac{1}{n} \sum_{i=1}^{n} \sum_{k=i}^{n} \binom{n}{k} (F(x))^k (1 - F(x))^{n-k} \\
&= \frac{1}{n} \sum_{k=1}^{n} \sum_{i=1}^{k} \binom{n}{k} (F(x))^k (1 - F(x))^{n-k} \\
&= \frac{1}{n} \sum_{k=1}^{n} k \binom{n}{k} (F(x))^k (1 - F(x))^{n-k} \\
&= \frac{1}{n} \left[n(F(x)) \right] = F(x),
\end{aligned}$$

so that (4.2.1) follows. □

The previous theorem claims that if you order a sample and randomly choose one of its elements, then you recover the original distribution. We capitalize on this idea in order to define a bivariate distribution with given marginals F and G, which was first formalized in Baker (2008) in the context of copula theory.

Definition 4.2 Fix $m, n \geq 1$. Denote by $\{F_{i:n}\}_{i=1}^{n}$ and $\{G_{i:m}\}_{i=1}^{m}$ the order statistics of F and G. Let $\boldsymbol{P} = \{p_{ij}\}_{i\in\{1,\ldots,n\},j\in\{1,\ldots,m\}}$ be a non–negative matrix such that

$$\boldsymbol{P}\boldsymbol{e} = \boldsymbol{e} \quad \text{and} \quad \boldsymbol{e}'\boldsymbol{P} = (n/m)\boldsymbol{e}'. \tag{4.2.2}$$

Define the bivariate distribution $H_{\boldsymbol{P}} : \mathbb{R}^2 \to [0,1]$ by

$$H_{\boldsymbol{P}}(x,y) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} F_{i:n}(x) G_{j:m}(y). \tag{4.2.3}$$

We call $H_{\boldsymbol{P}}$ a **bivariate distribution of the Baker–type**.

Theorem 4.3 *Let F and G be distribution functions over $\mathbb{R}$ with means μ and ν, and finite variances σ_F^2 and σ_G^2. Let $\{\mu_{i:n}\}_{i=1}^{n}$ and $\{\nu_{j:m}\}_{j=1}^{m}$ correspond to the means of $\{F_{i:n}\}_{i=1}^{n}$ and $\{G_{j:m}\}_{j=1}^{m}$, respectively. Let $\boldsymbol{P}$ and $H_{\boldsymbol{P}}$ be as in Definition 4.2. If $(X,Y) \sim H_{\boldsymbol{P}}$, then*

$$X \sim F, \quad Y \sim G, \quad \textit{and}$$

$$\rho_{\boldsymbol{P}} = \frac{n^{-1}\boldsymbol{\mu}\boldsymbol{P}\boldsymbol{\nu}' - \mu\nu}{\sigma_F\sigma_G} \tag{4.2.4}$$

where $\rho_{\boldsymbol{P}}$ corresponds to the Pearson's correlation coefficient of the bivariate distribution $H_{\boldsymbol{P}}$, and

$$\boldsymbol{\mu} = (\mu_{1:n}, \ldots, \mu_{n:n}) \quad \textit{and} \quad \boldsymbol{\nu} = (\nu_{1:m}, \ldots, \nu_{m:m}).$$

Proof. Let $(X,Y) \sim H_F$. Note that

$$\begin{aligned}
\lim_{y\to\infty} H_{\boldsymbol{P}}(x,y) &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} F_{i:n}(x) \left(\lim_{y\to\infty} G_{j:m}(y)\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\sum_{j=1}^{m} p_{ij}\right) F_{i:n}(x) \\
&= \frac{1}{n}\sum_{i=1}^{n} F_{i:n}(x), \quad \text{and}
\end{aligned}$$

$$\begin{aligned}\lim_{x\to\infty} H_{\boldsymbol{P}}(x,y) &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}\left(\lim_{x\to\infty} F_{i:n}(x)\right) G_{j:m}(y)\\ &= \frac{1}{n}\sum_{j=1}^{m}\left(\sum_{i=1}^{n} p_{ij}\right) G_{j:m}(y)\\ &= \frac{1}{n}\sum_{j=1}^{m}\left(\frac{n}{m}\right) G_{j:m}(y)\\ &= \frac{1}{m}\sum_{j=1}^{m} G_{i:m}(y),\end{aligned}$$

so that Theorem 4.1 implies that $X \sim F$ and $Y \sim G$.

Now,

$$\rho_{\boldsymbol{P}} = \frac{\int_R \int_R xy H_{\boldsymbol{P}}(\mathrm{d}x,\mathrm{d}y) - \mu\nu}{\sigma_F \sigma_G}, \tag{4.2.5}$$

where

$$\begin{aligned}\int_R \int_R xy H_{\boldsymbol{P}}(\mathrm{d}x,\mathrm{d}y) &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}\left(\int_{\mathbb{R}} xF_{i:n}(\mathrm{d}x)\right)\left(\int_{\mathbb{R}} yG_{j:m}(\mathrm{d}y)\right)\\ &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}\mu_{i:n}\nu_{j:m}\\ &= \frac{1}{n}\boldsymbol{\mu}\boldsymbol{P}\boldsymbol{\nu}',\end{aligned}$$

implying (4.2.4). □

Remark 12 *The idea behind a bivariate distribution of the Baker–type is the following. Simulate a sample of i.i.d.r.v.'s* $\{X_i\}_{i=1}^n$ *and a sample of i.i.d.r.v.'s* $\{Y_j\}_{j=1}^m$. *Order these samples into* $\{X_{i:n}\}_{i=1}^n$ *and* $\{Y_{j:m}\}_{j=1}^m$, *and pair each* $X_{i:n}$ *(*$i \in \{1,\ldots,n\}$*) with an element of* $\{Y_{j:m}\}_{j=1}^m$ *according to the probability vector* $\boldsymbol{p}_i := (p_{i1}, p_{i2}, \ldots, p_{im})$, *so that we have* n *pairs in total. It may happen that two or more pairs have contain the same element* $Y_{j:m}$. *Finally, select randomly one of those pairs. Thus, all the elements in* $\{X_{i:n}\}_{i=1}^n$ *(ditto* $\{Y_{i:m}\}_{i=1}^m$*) were equally likely to be chosen, however, the pairing of ordered elements suggests that the each pair is correlated in some way, depending on the choice of* $\boldsymbol{P}$. *Indeed, if* $m = n$ *and* $\boldsymbol{P} = \boldsymbol{I}$, *then we expect to have highly correlated pairs, since we paired big (small) elements of* $\{X_i\}_{i=1}^n$ *with big (small) elements of* $\{Y_i\}_{i=1}^n$.

From now on, we focus on bivariate distributions of the Baker–type with $m = n$: in this case the conditions (4.2.2) mean that $\boldsymbol{P}$ belongs to the set $\mathscr{D}$ of **doubly stochastic matrices of size** $n \times n$.

Definition 4.4 Let $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$. We write $\boldsymbol{a} \prec_{\mathrm{m}} \boldsymbol{b}$ if

$$\boldsymbol{a}\boldsymbol{e} = \boldsymbol{b}\boldsymbol{e}, \quad \text{and}$$

$$\sum_{i=1}^{k} a_i \leq \sum_{i=1}^{k} b_i \quad \text{for all} \quad k \in \{1, \ldots, n\}.$$

Remark 13 *Definition 4.4 is closely related to the concept of majorisation of vectors, which has a slightly more complex definition; see Marshall et al. (2011) for more details on majorisation.*

Lemma 4.5 *Let $\boldsymbol{a} \in \mathbb{R}^n$ be such that $a_1 \leq a_2 \leq \cdots \leq a_n$. Then:*

1. *For any $\boldsymbol{P} \in \mathscr{D}$*
$$\boldsymbol{a} \prec_{\mathrm{m}} \boldsymbol{a}\boldsymbol{P}, \quad \textit{and} \tag{4.2.6}$$
2. *For any $\boldsymbol{P} \in \mathscr{D}$*
$$-\boldsymbol{a}\boldsymbol{I}^* \prec_{\mathrm{m}} -\boldsymbol{a}\boldsymbol{P},$$
where
$$\boldsymbol{I}^* = \begin{pmatrix} 0 & \cdots & 0 & 0 & 1 \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & 1 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 & 0 \end{pmatrix}.$$

Proof.

1. Fix $\boldsymbol{P} \in \mathscr{D}$, fix $k \in \{1, \ldots, n\}$ and define $t_i = \sum_{j=1}^{k} p_{ij}$ for all $i \in \{1, \ldots, n\}$. Then,
$$\sum_{j=1}^{k} (\boldsymbol{a}\boldsymbol{P})_j = \sum_{j=1}^{k} \sum_{i=1}^{n} a_i p_{ij} = \sum_{i=1}^{n} a_i t_i,$$
$$\sum_{i=1}^{n} t_i = \sum_{j=1}^{k} \left(\sum_{i=1}^{n} p_{ij}^* \right) = k, \quad \text{and}$$

$$\begin{aligned}\sum_{j=1}^{k}(\boldsymbol{aP})_j - \sum_{j=1}^{k} a_j &= \sum_{i=1}^{n} a_i t_i - \sum_{j=1}^{k} a_j \\ &= \sum_{i=1}^{n} a_i t_i - \sum_{j=1}^{k} a_j + a_k \left(k - \sum_{i=1}^{n} t_i\right) \\ &= \sum_{i=1}^{k}(a_i - a_k)(t_i - 1) + \sum_{i=k+1}^{n} t_i(a_i - a_k) \geq 0,\end{aligned}$$

so that $\sum_{j=1}^{k} a_j \leq \sum_{j=1}^{k}(\boldsymbol{aP})_j$. The fact that $(\boldsymbol{aP})\boldsymbol{e} = \boldsymbol{a}(\boldsymbol{Pe}) = \boldsymbol{ae}$ proves that $\boldsymbol{a} \prec_{\mathrm{m}} \boldsymbol{aP}$.

2. Fix $\boldsymbol{P} \in \mathscr{D}$. The vector $\boldsymbol{b} := -a\boldsymbol{I}^*$ has the form $\boldsymbol{b} = (-a_n, -a_{n-1}, \ldots, -a_1)$ and thus, $b_1 \leq b_2 \leq \cdots \leq b_n$. Let $\boldsymbol{P}^* := \boldsymbol{I}^*\boldsymbol{P}$, which is itself an element of $\mathscr{D}$. Then, according to (4.2.6),

$$\boldsymbol{b} \prec_{\mathrm{m}} \boldsymbol{bP}^*,$$

or in other words,

$$-\boldsymbol{aI}^* \prec_{\mathrm{m}} (-\boldsymbol{aI}^*)(\boldsymbol{I}^*\boldsymbol{P}) = -\boldsymbol{aP}$$

and the proof is finished.

□

Lemma 4.6 *Let $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^n$ where $\boldsymbol{a} \prec_{\mathrm{m}} \boldsymbol{b}$ and $\boldsymbol{c}$ is such that $c_1 \leq c_2 \leq \cdots \leq c_n$. Then $\boldsymbol{ac}' \geq \boldsymbol{bc}'$.*

Proof. Let $d_1 = c_1, d_j = c_j - c_{j-1}, j \in \{2, \ldots, n\}$, so that $d_j \geq 0$ for all $j \in \{2, \ldots, n\}$. Notice that $\boldsymbol{a} \prec_{\mathrm{m}} \boldsymbol{b}$ implies that

$$\sum_{i=j}^{n} a_i \geq \sum_{i=j}^{n} b_i \quad \text{for all} \quad j \in \{2, \ldots, n\}$$

Then,

$$\begin{aligned}
\boldsymbol{ac}' &= \sum_{i=1}^{n} a_i c_i = \sum_{i=1}^{n} a_i \left(\sum_{j=1}^{i} d_j \right) = \sum_{j=1}^{n} d_j \sum_{i=j}^{n} a_i \\
&= d_1 \sum_{i=1}^{n} a_i + \sum_{j=2}^{n} d_j \left(\sum_{i=j}^{n} a_i \right) = d_1 \sum_{i=1}^{n} b_i + \sum_{j=2}^{n} d_j \left(\sum_{i=j}^{n} a_i \right) \\
&\geq d_1 \sum_{i=1}^{n} b_i + \sum_{j=2}^{n} d_j \left(\sum_{i=j}^{n} b_i \right) = \sum_{j=1}^{n} d_j \sum_{i=j}^{n} b_i \\
&= \sum_{i=1}^{n} b_i \left(\sum_{j=1}^{i} d_j \right) = \sum_{i=1}^{n} b_i c_i = \boldsymbol{bc}'.
\end{aligned}$$

□

Theorem 4.7 *Let $\boldsymbol{P} \in \mathscr{D}$ and let $\rho_{\boldsymbol{P}}$ be the Pearson's correlation coefficient of the bivariate distribution $H_{\boldsymbol{P}}$ defined in (4.2.3). Then,*

$$\frac{n^{-1}\boldsymbol{\mu I}^*\boldsymbol{\nu}' - \mu\nu}{\sigma_F \sigma_G} \leq \rho_{\boldsymbol{P}} \leq \frac{n^{-1}\boldsymbol{\mu\nu}' - \mu\nu}{\sigma_F \sigma_G}, \tag{4.2.7}$$

where the lower bound is attained if $\boldsymbol{P} = \boldsymbol{I}^*$ and the upper bound is attained if $\boldsymbol{P} = \boldsymbol{I}$.

Proof. From Theorem 4.3 we know that $\rho_{\boldsymbol{P}} = \frac{n^{-1}\boldsymbol{\mu P\nu}' - \mu\nu}{\sigma_F \sigma_G}$. Lemma 4.5 (Item 1) and Lemma 4.5 imply that $\boldsymbol{\mu P\nu}' \leq \boldsymbol{\mu\nu}'$, so the upper bound in (4.2.7) follows and is attained if $\boldsymbol{P} = \boldsymbol{I}$. Lemma 4.5 (Item 2) and Lemma 4.5 imply that $-\boldsymbol{\mu P\nu}' \leq -\boldsymbol{\mu I}^*\boldsymbol{\nu}'$, so the lower bound in (4.2.7) follows and is attained if $\boldsymbol{P} = \boldsymbol{I}^*$. □

Remark 14 *Theorem 4.7 formalizes what was outlined in Remark 12: the Baker–type bivariate distribution which results from pairing big (small) observations of $\{X_i\}_{i=n}$ with big (small) observations of $\{Y_i\}_{i=n}$ has the highest Pearson's correlation possible amongst the class of Baker–type bivariate distributions based on n order statistics. On the other hand, we also get that the Baker–type bivariate distribution which results from pairing big (small) observations of $\{X_i\}_{i=n}$ with small (big) observations of $\{Y_i\}_{i=n}$ has the lowest Pearson's correlation possible amongst the class of Baker–type bivariate distributions based on n order statistics.*

We will denote by $(H_+^{(n)}, \rho_+^{(n)})$ the pair $(H_{\boldsymbol{I}}, \rho_{\boldsymbol{I}})$, and by $(H_-^{(n)}, \rho_-^{(n)})$ the pair $(H_{\boldsymbol{I}^*}, \rho_{\boldsymbol{I}^*})$. In Dou et al. (2013) it is proved that

$$H_+^{(n)} \xrightarrow{w} F \wedge G \quad \text{and} \quad H_-^{(n)} \xrightarrow{w} (F+G-1) \vee 0 \quad \text{as} \quad n \to \infty,$$

where $F \wedge G$ is the **Frechét–Hoeffding upper bound** and $(F+G-1)\vee 0$ is the **Frechét–Hoeffding lower bound**. Furthermore, it is proved that

$$\rho_+^{(n)} \to \rho_{\max} \quad \text{and} \quad \rho_-^{(n)} \to \rho_{\min} \quad \text{as} \quad n \to \infty,$$

where $\rho_{\max}$ corresponds to the Pearson's correlation coeficient of $F \wedge G$ and $\rho_{\min}$ corresponds to the Pearson's correlation coeficient of $(F+G-1)\vee 0$. $\rho_{\max}$ and $\rho_{\min}$ correspond to the maximal and minimal Pearson's correlation coefficient that a bivariate distribution with marginals F and G can have.

Lemma 4.8 *Let F and G be distributions with finite variances and let $\rho_0 \in (\rho_{\min}, \rho_{\max})$. Then, there exits $n_0 \geq 1$ and a doubly stochastic matrix $\boldsymbol{P}_0$ of size $n_0 \times n_0$ such that $H_{\boldsymbol{P}_0}$ is a bivariate distribution with marginals F and G, with Pearson's correlation coefficient $\rho_{\boldsymbol{P}} = \rho_0$. Moreover, if $n_0 \geq 2$ then $\boldsymbol{P}_0$ is on the form*

$$\boldsymbol{P}_0 = \begin{cases} \frac{\rho_0}{\rho_+^{(n_0)}} \boldsymbol{I} + \left(1 - \frac{\rho_0}{\rho_+^{(n_0)}}\right)(n_0^{-1}\boldsymbol{e}\boldsymbol{e}') & \text{if} \quad \rho_0 \in [0, \rho_{\max}) \\ \frac{\rho_0}{\rho_-^{(n_0)}} \boldsymbol{I}^* + \left(1 - \frac{\rho_0}{\rho_-^{(n_0)}}\right)(n_0^{-1}\boldsymbol{e}\boldsymbol{e}') & \text{if} \quad \rho_0 \in (\rho_{\min}, 0) \end{cases} \tag{4.2.8}$$

Proof. Suppose that $\rho_0 \in [0, \rho_{\max})$. Since $\rho_+^{(n)} \to \rho_{\max}$, then there exists $n_0 \geq 1$ such that $\rho_+^{(n_0)} \geq \rho_0$. In this case, $\boldsymbol{P}_0$ defined as in (4.2.8) is indeed doubly stochastic, and

$$\begin{aligned} \rho_{\boldsymbol{P}_0} &= \frac{n_0^{-1}\boldsymbol{\mu}\boldsymbol{P}_0\boldsymbol{\nu}' - \mu\nu}{\sigma_F\sigma_G} \\ &= \frac{n_0^{-1}\boldsymbol{\mu}\left[\frac{\rho_0}{\rho_+^{(n_0)}}\boldsymbol{I} + \left(1 - \frac{\rho_0}{\rho_+^{(n_0)}}\right)(n_0^{-1}\boldsymbol{e}\boldsymbol{e}')\right]\boldsymbol{\nu}' - \mu\nu}{\sigma_F\sigma_G} \\ &= \frac{\frac{\rho_0}{\rho_+^{(n_0)}} n_0^{-1}\boldsymbol{\mu}\boldsymbol{\nu} + \left(1 - \frac{\rho_0}{\rho_+^{(n_0)}}\right)(n_0^{-1}\boldsymbol{\mu}\boldsymbol{e})(n_0^{-1}\boldsymbol{e}'\boldsymbol{\nu}') - \mu\nu}{\sigma_F\sigma_G} \\ &= \frac{\frac{\rho_0}{\rho_+^{(n_0)}} n_0^{-1}\boldsymbol{\mu}\boldsymbol{\nu} + \left(1 - \frac{\rho_0}{\rho_+^{(n_0)}}\right)\mu\nu - \mu\nu}{\sigma_F\sigma_G} \\ &= \frac{\rho_0}{\rho_+^{(n_0)}}\left\{\frac{n_0^{-1}\boldsymbol{\mu}\boldsymbol{\nu}' - \mu\nu}{\sigma_F\sigma_G}\right\} = \frac{\rho_0}{\rho_+^{(n_0)}}\rho_+^{(n_0)} = \rho_0. \end{aligned}$$

Now suppose that $\rho_0 \in (\rho_{\max}, n_0)$. Since $\rho_-^{(n)} \to \rho_{\min}$, then there exists $n_0 \geq 1$ such that $\rho_-^{(n_0)} \leq \rho_0$, which guarantees that $\boldsymbol{P}_0$ defined as in (4.2.8) is indeed doubly stochastic. Then,

$$\begin{aligned}
\rho_{\boldsymbol{P}_0} &= \frac{n_0^{-1}\boldsymbol{\mu}\boldsymbol{P}_0\boldsymbol{\nu}' - \mu\nu}{\sigma_F\sigma_G} \\
&= \frac{n_0^{-1}\boldsymbol{\mu}\left[\frac{\rho_0}{\rho_-^{(n_0)}}\boldsymbol{I}^* + \left(1 - \frac{\rho_0}{\rho_-^{(n_0)}}\right)(n_0^{-1}\boldsymbol{e}\boldsymbol{e}')\right]\boldsymbol{\nu}' - \mu\nu}{\sigma_F\sigma_G} \\
&= \frac{\frac{\rho_0}{\rho_-^{(n_0)}}n_0^{-1}\boldsymbol{\mu}\boldsymbol{I}^*\boldsymbol{\nu} + \left(1 - \frac{\rho_0}{\rho_-^{(n_0)}}\right)(n_0^{-1}\boldsymbol{\mu}\boldsymbol{e})(n_0^{-1}\boldsymbol{e}'\boldsymbol{\nu}') - \mu\nu}{\sigma_F\sigma_G} \\
&= \frac{\frac{\rho_0}{\rho_-^{(n_0)}}n_0^{-1}\boldsymbol{\mu}\boldsymbol{I}^*\boldsymbol{\nu} + \left(1 - \frac{\rho_0}{\rho_-^{(n_0)}}\right)\mu\nu - \mu\nu}{\sigma_F\sigma_G} \\
&= \frac{\rho_0}{\rho_-^{(n_0)}}\left\{\frac{n_0^{-1}\boldsymbol{\mu}\boldsymbol{I}^*\boldsymbol{\nu}' - \mu\nu}{\sigma_F\sigma_G}\right\} = \frac{\rho_0}{\rho_-^{(n_0)}}\rho_-^{(n_0)} = \rho_0,
\end{aligned}$$

and the proof is finished. Lemma 4.8 provides an explicit way to construct bivariate distributions with given marginals and Pearson's correlation coefficient based on the distribution of the order statistics of said marginals.

4.3 Bivariate phase–type distributions of the Baker–type

In the following we give a matrix representation of the Baker–type bivariate distirbutions with phase–type marginals. In particular, we show that they belong to the class MPH* of Kulkarni's multivariate phase–type distributions.

Theorem 4.9 *Let F and G be phase–type–distributed with parameters $(\boldsymbol{\alpha}, \boldsymbol{S})$ and $(\boldsymbol{\pi}, \boldsymbol{T})$, respectively. Let $\boldsymbol{P}$ be a doubly stochastic matrix of size $n \times n$, and let $H_{\boldsymbol{P}}$ be the bivariate distribution of the Baker–type defined by (4.2.3). Then, $H_{\boldsymbol{P}}$ corresponds to*

$$\text{MPH}^*\left((\boldsymbol{\alpha}^{(n)}, \boldsymbol{0}), \begin{pmatrix} \boldsymbol{S}^{(n)} & \boldsymbol{D}_{\boldsymbol{P}} \\ \boldsymbol{0} & \boldsymbol{T}^{(n)} \end{pmatrix}, \begin{pmatrix} \boldsymbol{e} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{e} \end{pmatrix}\right), \tag{4.3.1}$$

where

$$\boldsymbol{\alpha}^{(n)} = n^{-1}(\boldsymbol{\alpha}_{1:n}, \boldsymbol{\alpha}_{2:n}, \ldots, \boldsymbol{\alpha}_{n:n}), \quad \boldsymbol{S}^{(n)} = \begin{pmatrix} \boldsymbol{S}_{1:n} & 0 & \cdots & 0 \\ \mathbf{0} & \boldsymbol{S}_{2:n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{S}_{n:n} \end{pmatrix},$$

$$\boldsymbol{T}^{(n)} = \begin{pmatrix} \boldsymbol{T}_{1:n} & 0 & \cdots & 0 \\ \mathbf{0} & \boldsymbol{T}_{2:n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{T}_{n:n} \end{pmatrix}, \quad \textit{and}$$

$$\boldsymbol{D}_{\boldsymbol{P}} = \begin{pmatrix} p_{11}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{1:n} & p_{12}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{1n}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{n:n} \\ p_{21}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{1:n} & p_{22}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{2n}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{n:n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{1:n} & p_{n2}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{nn}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{n:n} \end{pmatrix}.$$

Proof. Let $h : \mathbb{R}_+^2 \to \mathbb{R}_+$ be the density function of (4.3.1) and let $(X, Y) \sim h$. Let $\{J_t\}_{t\geq 0}$ be the underlying Markov jump process of h with state space $\mathfrak{E}_S \cup \mathfrak{E}_T$. $\{J_t\}_{t\geq 0}$ starts in some state in $\mathfrak{E}_S$ according to the probability vector $\boldsymbol{\alpha}^{(n)}$, evolves within $\mathfrak{E}_S$ according to $\boldsymbol{S}^{(n)}$, jumps to $\mathfrak{E}_T$ with intensity $\boldsymbol{D}_{\boldsymbol{P}}$, evolves within $\mathfrak{E}_T$ according to $\boldsymbol{T}^{(n)}$, and gets absorbed from there. The rewards in (4.3.1) imply that X and Y correspond to the occupation times of $\{J_t\}_{t\geq 0}$ in $\mathfrak{E}_S$ and $\mathfrak{E}_T$. Thus, h is on the form

$$h(x, y) = \boldsymbol{\alpha}^{(n)} e^{\boldsymbol{S}^{(n)}x} \boldsymbol{D}_{\boldsymbol{P}} e^{\boldsymbol{T}^{(n)}x} \boldsymbol{t}^{(n)}, \quad x, y \geq 0, \tag{4.3.2}$$

with

$$\boldsymbol{t}^{(n)} = -\boldsymbol{T}^{(n)}\boldsymbol{e} = \begin{pmatrix} \boldsymbol{t}_{1:n} \\ \boldsymbol{t}_{2:n} \\ \vdots \\ \boldsymbol{t}_{n:n} \end{pmatrix}.$$

Now, for $x, y \geq 0$,

$$
\begin{aligned}
h(x,y) &= \frac{1}{n}(\boldsymbol{\alpha}_{1:n}, \boldsymbol{\alpha}_{2:n}, \ldots, \boldsymbol{\alpha}_{n:n}) \begin{pmatrix} e^{\boldsymbol{S}_{1:n}x} & 0 & \cdots & 0 \\ \mathbf{0} & e^{\boldsymbol{S}_{2:n}x} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & e^{\boldsymbol{S}_{n:n}x} \end{pmatrix} \\
&\quad \times \boldsymbol{D}_{\boldsymbol{P}} \begin{pmatrix} e^{\boldsymbol{T}_{1:n}y} & 0 & \cdots & 0 \\ \mathbf{0} & e^{\boldsymbol{T}_{2:n}y} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & e^{\boldsymbol{T}_{n:n}y} \end{pmatrix} \begin{pmatrix} \boldsymbol{t}_{1:n} \\ \boldsymbol{t}_{2:n} \\ \vdots \\ \boldsymbol{t}_{n:n} \end{pmatrix} \\
&= \frac{1}{n}(\boldsymbol{\alpha}_{1:n}e^{\boldsymbol{S}_{1:n}x}, \boldsymbol{\alpha}_{2:n}e^{\boldsymbol{S}_{2:n}x}, \ldots, \boldsymbol{\alpha}_{n:n}e^{\boldsymbol{S}_{n:n}x}) \\
&\quad \times \begin{pmatrix} p_{11}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{1:n} & p_{12}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{1n}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{n:n} \\ p_{21}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{1:n} & p_{22}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{2n}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{n:n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{1:n} & p_{n2}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{nn}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{n:n} \end{pmatrix} \begin{pmatrix} e^{\boldsymbol{T}_{1:n}y}\boldsymbol{t}_{1:n} \\ e^{\boldsymbol{T}_{2:n}y}\boldsymbol{t}_{2:n} \\ \vdots \\ e^{\boldsymbol{T}_{n:n}y}\boldsymbol{t}_{n:n} \end{pmatrix} \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}\left(\boldsymbol{\alpha}_{i:n}e^{\boldsymbol{S}_{i:n}x}\boldsymbol{s}_{i:n}\right)\left(\boldsymbol{\pi}_{j:n}e^{\boldsymbol{T}_{j:n}y}\boldsymbol{t}_{j:n}\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij} f_{i:n}(x) g_{j:n}(y),
\end{aligned}
$$

where $f_{i:n}$ and $g_{j:n}$ correspond to the densities of $F_{i:n}$ and $G_{j:n}$, respectively. This means that h is the density of the bivariate distribution $H_{\boldsymbol{P}}$ of the Baker–type defined in (4.2.3), and the proof is finished. □

Remark 15 *In Bladt and Nielsen (2010b), a bivariate exponential distribution in* MPH* *with arbitrary correlation coefficient was constructed. This inspired the extension to bivariate phase–type distributions with arbitrary correlation coefficient constructed in Theorem 4.9. The construction in Bladt and Nielsen (2010b) lays on an implicit Baker–type bivariate distribution, however, time–reversal arguments were used to attain low dimensionality of the parameters. Years later, He et al. (2012) characterised the bivariate exponential distribution with the highest correlation coefficient within the class of bivariate exponential distributions of the type constructed in Bladt and Nielsen (2010b) with fixed dimension: in our setting, their result follows directly from Theorem 4.7 and Theorem 4.9.*

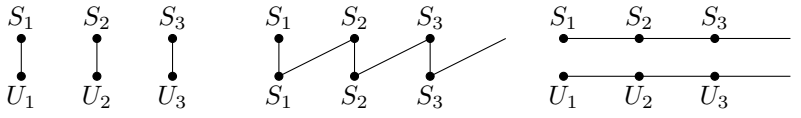
4.4 Dependent Sparre–Andersen processes

In the following we define three classes of dependent Sparre–Andersen processes. The terminology "Sparre–Andersen" will allude to the fact that on the three classes, $\{R_t\}_{t\geq 0}$ is on the form

$$R_t = u + pt - \sum_{i\geq 1: t\leq T_i} U_i,$$

where the interarrival times $\{S_i\}_{i\geq 1}$ ($S_i = T_i - T_{i-1}$) are identically distributed with common distribution F, and the claim sizes $\{U_i\}_{i\geq 1}$ are identically distributed with common distribution G. The three cases of dependence we consider are:

1. **By pairs.** There exists a fixed Baker–type dependence between S_i and U_i for all $i \geq 0$. The bivariate sequence $\{(S_i, U_i)\}_{i\geq 1}$ is i.i.d..
2. **Alternating.** There exists a fixed Baker–type dependence between S_i and U_i for all $i \geq 1$, and a fixed Baker–type dependence between U_i and S_{i+1} for all $i \geq 1$. The alternating nature of this dependence imply that are no pair of independent elements in the set $\{\{S_i\}_{i\geq 1}, \{U_i\}_{i\geq 1}\}$: everything is dependent on everything else.
3. **Sequential.** The sequence $\{S_i\}_{i\geq 1}$ is independent from the sequence $\{U_i\}_{i\geq 1}$. There exists a fixed Baker–type dependence between S_{i-1} and S_i for all $i \geq 1$, and a fixed Baker–type dependence between U_i and U_{i+1} for all $i \geq 1$. Thus, any given S_i is independent from any given U_j, but there is no pair of independent elements in $\{S_i\}_{i\geq 1}$ or in $\{U_i\}_{i\geq 1}$.



(a) By pairs **(b)** Alternating **(c)** Sequential

Figure 4.1: (a) Dependence by pairs: U_i is drawn conditionally on S_i; (b) Alternating dependence: U_i is drawn conditionally on S_i, and S_{i+1} is drawn conditionally on U_i; (c) Sequential dependence: Independence between $\{S_i\}_{i\geq 1}$ and $\{U_i\}_{i\geq 1}$ while dependence within the elements of each sequence.

If $\boldsymbol{F}$ and $\boldsymbol{G}$ correspond to PH$(\boldsymbol{\alpha}, \boldsymbol{S})$ and PH$(\boldsymbol{\pi}, \boldsymbol{T})$, respectively, we can regard the three cases of dependent Sparre–Andersen processes as fluid flow risk process as follows.

4.4.1 By pairs

Let $\boldsymbol{P} = \{p_{ij}\}$ be a $n \times n$ doubly stochastic matrix and let

$$\boldsymbol{\alpha}^{(n)} = n^{-1}(\boldsymbol{\alpha}_{1:n}, \boldsymbol{\alpha}_{2:n}, \ldots, \boldsymbol{\alpha}_{n:n}), \quad \boldsymbol{S}^{(n)} = \begin{pmatrix} \boldsymbol{S}_{1:n} & 0 & \cdots & 0 \\ \boldsymbol{0} & \boldsymbol{S}_{2:n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{S}_{n:n} \end{pmatrix},$$

$$\boldsymbol{T}^{(n)} = \begin{pmatrix} \boldsymbol{T}_{1:n} & 0 & \cdots & 0 \\ \boldsymbol{0} & \boldsymbol{T}_{2:n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{T}_{n:n} \end{pmatrix}, \quad \boldsymbol{t}^{(n)} = -\boldsymbol{T}^{(n)}\boldsymbol{e}, \quad \text{and}$$

$$\boldsymbol{D}_{\boldsymbol{P}} = \begin{pmatrix} p_{11}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{1:n} & p_{12}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{1n}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{n:n} \\ p_{21}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{1:n} & p_{22}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{2n}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{n:n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{1:n} & p_{n2}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{nn}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{n:n} \end{pmatrix}.$$

Let $\mathfrak{E}_{S^{(n)}}$ and $\mathfrak{E}_{T^{(n)}}$ be the phase–spaces associated to $\boldsymbol{S}^{(n)}$ and $\boldsymbol{T}^{(n)}$, respectively. Let $\{R_t\}_{t\geq 0}$ be the fluid flow risk process with associated fluid flow process $\{(V_t, J_t)\}_{t\geq 0}$, where $\{J_t\}_{t\geq 0}$ has state–space $\mathfrak{E}_+ \cup \mathfrak{E}_-$ with $\mathfrak{E}_+ = \mathfrak{E}_{S^{(n)}}$ and $\mathfrak{E}_- = \mathfrak{E}_{T^{(n)}}$, $J_0 \sim (\boldsymbol{\alpha}^{(n)}, \boldsymbol{0})$ and has an intensity matrix given by

$$\begin{pmatrix} \boldsymbol{S}^{(n)} & \boldsymbol{D}_{\boldsymbol{P}} \\ \boldsymbol{t}^{(n)}\boldsymbol{\alpha}^{(n)} & \boldsymbol{T}^{(n)} \end{pmatrix}. \tag{4.4.1}$$

Let the reward rates $\{r_i\}_{i\in\mathfrak{E}_+\cup\mathfrak{E}_-}$ and Brownian noise rates $\{\sigma_i\}_{i\in\mathfrak{E}_+\cup\mathfrak{E}_-}$ be given by

$$\begin{aligned} r_i = p \quad &\text{and} \quad \sigma_i = 0, \quad i \in \mathfrak{E}_+, \\ r_i = -1 \quad &\text{and} \quad \sigma_i = 0, \quad i \in \mathfrak{E}_-. \end{aligned}$$

We claim that $\{R_t\}_{t\geq 0}$ corresponds to a Sparre–Andersen process with dependence by pairs. Indeed, the structure of (4.4.1) implies that

$$\begin{aligned} &\mathbb{P}(S_1 \in \mathrm{d}s_1, U_1 \in \mathrm{d}u_1, S_2 \in \mathrm{d}s_2, U_2 \in \mathrm{d}u_2, \ldots, S_k \in \mathrm{d}s_k, U_k \in \mathrm{d}u_k) \\ &= \boldsymbol{\alpha}^{(n)} \left(\prod_{i=1}^{k} e^{\boldsymbol{S}^{(n)}s_1}\boldsymbol{D}_{\boldsymbol{P}}e^{\boldsymbol{T}^{(n)}u_1} \left(\boldsymbol{t}^{(n)}\boldsymbol{\alpha}^{(n)}\right) \right) \boldsymbol{e}\mathrm{d}s_1\mathrm{d}u_1\mathrm{d}s_2\mathrm{d}u_2 \ldots \mathrm{d}s_k\mathrm{d}u_k \\ &= \prod_{i=1}^{k} \left(\boldsymbol{\alpha}^{(n)}e^{\boldsymbol{S}^{(n)}s_i}\boldsymbol{D}_{\boldsymbol{P}}e^{\boldsymbol{T}^{(n)}u_i}\boldsymbol{t}^{(n)}\mathrm{d}s_i\mathrm{d}u_i \right), \end{aligned}$$

so that

$$\mathbb{P}(S_1 \in \mathrm{d}s_k, U_1 \in \mathrm{d}u_k) = \boldsymbol{\alpha}^{(n)} e^{\boldsymbol{S}^{(n)} s_k} \boldsymbol{D_P} e^{\boldsymbol{T}^{(n)} u_k} \boldsymbol{t}^{(n)} \mathrm{d}s_k \mathrm{d}u_k, \quad k \geq 1,$$

which according to Theorem 4.9 corresponds to the Baker–type distribution

$$H_{\boldsymbol{P}}(x,y) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} F_{i:n}(x) G_{j:m}(y).$$

Moreover, $\{(S_i, U_i)\}_{i\geq 0}$ is a sequence of independent random vectors.

4.4.2 Alternating

Let $\boldsymbol{P} = \{p_{ij}\}$ and $\boldsymbol{Q} = \{q_{ij}\}$ be $n \times n$ doubly stochastic matrices and let

$$\boldsymbol{\alpha}^{(n)} = n^{-1}(\boldsymbol{\alpha}_{1:n}, \boldsymbol{\alpha}_{2:n}, \ldots, \boldsymbol{\alpha}_{n:n}), \quad \boldsymbol{S}^{(n)} = \begin{pmatrix} \boldsymbol{S}_{1:n} & 0 & \cdots & 0 \\ \boldsymbol{0} & \boldsymbol{S}_{2:n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{S}_{n:n} \end{pmatrix},$$

$$\boldsymbol{\pi}^{(n)} = n^{-1}(\boldsymbol{\pi}_{1:n}, \boldsymbol{\pi}_{2:n}, \ldots, \boldsymbol{\pi}_{n:n}), \quad \boldsymbol{T}^{(n)} = \begin{pmatrix} \boldsymbol{T}_{1:n} & 0 & \cdots & 0 \\ \boldsymbol{0} & \boldsymbol{T}_{2:n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{T}_{n:n} \end{pmatrix},$$

$$\boldsymbol{D_P} = \begin{pmatrix} p_{11}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{1:n} & p_{12}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{1n}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{n:n} \\ p_{21}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{1:n} & p_{22}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{2n}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{n:n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{1:n} & p_{n2}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{nn}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{n:n} \end{pmatrix}, \quad \text{and}$$

$$\boldsymbol{E_Q} = \begin{pmatrix} q_{11}\boldsymbol{t}_{1:n}\boldsymbol{\alpha}_{1:n} & q_{12}\boldsymbol{t}_{1:n}\boldsymbol{\alpha}_{2:n} & \cdots & q_{1n}\boldsymbol{t}_{1:n}\boldsymbol{\alpha}_{n:n} \\ q_{21}\boldsymbol{t}_{2:n}\boldsymbol{\alpha}_{1:n} & q_{22}\boldsymbol{t}_{2:n}\boldsymbol{\alpha}_{2:n} & \cdots & q_{2n}\boldsymbol{t}_{2:n}\boldsymbol{\alpha}_{n:n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1}\boldsymbol{t}_{n:n}\boldsymbol{\alpha}_{1:n} & q_{n2}\boldsymbol{t}_{n:n}\boldsymbol{\alpha}_{2:n} & \cdots & q_{nn}\boldsymbol{t}_{n:n}\boldsymbol{\alpha}_{n:n} \end{pmatrix}.$$

Let $\mathfrak{E}_{S^{(n)}}$ and $\mathfrak{E}_{T^{(n)}}$ be the phase–spaces associated to $\boldsymbol{S}^{(n)}$ and $\boldsymbol{T}^{(n)}$, respectively. Let $\{R_t\}_{t\geq 0}$ be the fluid flow risk process with associated fluid flow process $\{(V_t, J_t)\}_{t\geq 0}$, where $\{J_t\}_{t\geq 0}$ has state–space $\mathfrak{E}_+ \cup \mathfrak{E}_-$ with $\mathfrak{E}_+ = \mathfrak{E}_{S^{(n)}}$ and $\mathfrak{E}_- = \mathfrak{E}_{T^{(n)}}$, $J_0 \sim (\boldsymbol{\alpha}^{(n)}, \boldsymbol{0})$ and has an intensity matrix given by

$$\begin{pmatrix} \boldsymbol{S}^{(n)} & \boldsymbol{D_P} \\ \boldsymbol{E_Q} & \boldsymbol{T}^{(n)} \end{pmatrix}. \tag{4.4.2}$$

Let the reward rates $\{r_i\}_{i\in\mathfrak{E}_+\cup\mathfrak{E}_-}$ and Brownian noise rates $\{\sigma_i\}_{i\in\mathfrak{E}_+\cup\mathfrak{E}_-}$ be given by

$$\begin{aligned} r_i = p \quad &\text{and} \quad \sigma_i = 0, \quad i \in \mathfrak{E}_+, \\ r_i = -1 \quad &\text{and} \quad \sigma_i = 0, \quad i \in \mathfrak{E}_-. \end{aligned}$$

We claim that $\{R_t\}_{t\geq 0}$ corresponds to a Sparre–Andersen process with alternating dependence. Indeed, the structure of (4.4.8) implies that

$$\begin{aligned} &\mathbb{P}(S_1 \in \mathrm{d}s_1, U_1 \in \mathrm{d}u_1, S_2 \in \mathrm{d}s_2, U_2 \in \mathrm{d}u_2, \ldots, S_k \in \mathrm{d}s_k, U_k \in \mathrm{d}u_k, S_{k+1} \in \mathrm{d}s_{k+1}) \\ &= \boldsymbol{\alpha}^{(n)} \left(\prod_{i=1}^{k} e^{\boldsymbol{S}^{(n)} s_i} \boldsymbol{D}_{\boldsymbol{P}} e^{\boldsymbol{T}^{(n)} u_i} \boldsymbol{E}_{\boldsymbol{Q}} \right) \boldsymbol{e} \mathrm{d}s_1 \mathrm{d}u_1 \mathrm{d}s_2 \mathrm{d}u_2 \ldots \mathrm{d}s_k \mathrm{d}u_k \mathrm{d}s_{k+1}. \end{aligned}$$

Integrating with respect to $s_1, \ldots, s_{k-1}, u_1, \ldots, u_{k-1}$ we get that

$$\begin{aligned} &\mathbb{P}(S_k \in \mathrm{d}s_k, U_k \in \mathrm{d}u_k, S_{k+1} \in \mathrm{d}s_{k+1}) \\ &= \boldsymbol{\alpha}^{(n)} \left(\prod_{i=1}^{k} (-\boldsymbol{S}^{(n)})^{-1} \boldsymbol{D}_{\boldsymbol{P}} (-\boldsymbol{T}^{(n)})^{-1} \boldsymbol{E}_{\boldsymbol{Q}} \right) \qquad (4.4.3) \\ &\quad \times e^{\boldsymbol{S}^{(n)} s_k} \boldsymbol{D}_{\boldsymbol{P}} e^{\boldsymbol{T}^{(n)} u_k} \boldsymbol{E}_{\boldsymbol{Q}} e^{\boldsymbol{S}^{(n)} s_{k+1}} \boldsymbol{D}_{\boldsymbol{P}} \boldsymbol{e} \mathrm{d}s_k \mathrm{d}u_k \mathrm{d}s_{k+1}. \qquad (4.4.4) \end{aligned}$$

Now, notice that

$$\begin{aligned} (-\boldsymbol{S}^{(n)})^{-1} \boldsymbol{D}_{\boldsymbol{P}} &= \begin{pmatrix} (-\boldsymbol{S}_{1:n})^{-1} & 0 & \cdots & 0 \\ \boldsymbol{0} & (-\boldsymbol{S}_{2:n})^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & (-\boldsymbol{S}_{n:n})^{-1} \end{pmatrix} \\ &\quad \times \begin{pmatrix} p_{11} \boldsymbol{s}_{1:n} \boldsymbol{\pi}_{1:n} & p_{12} \boldsymbol{s}_{1:n} \boldsymbol{\pi}_{2:n} & \cdots & p_{1n} \boldsymbol{s}_{1:n} \boldsymbol{\pi}_{n:n} \\ p_{21} \boldsymbol{s}_{2:n} \boldsymbol{\pi}_{1:n} & p_{22} \boldsymbol{s}_{2:n} \boldsymbol{\pi}_{2:n} & \cdots & p_{2n} \boldsymbol{s}_{2:n} \boldsymbol{\pi}_{n:n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} \boldsymbol{s}_{n:n} \boldsymbol{\pi}_{1:n} & p_{n2} \boldsymbol{s}_{n:n} \boldsymbol{\pi}_{2:n} & \cdots & p_{nn} \boldsymbol{s}_{n:n} \boldsymbol{\pi}_{n:n} \end{pmatrix} \\ &= \begin{pmatrix} p_{11} \boldsymbol{e} \boldsymbol{\pi}_{1:n} & p_{12} \boldsymbol{e} \boldsymbol{\pi}_{2:n} & \cdots & p_{1n} \boldsymbol{e} \boldsymbol{\pi}_{n:n} \\ p_{21} \boldsymbol{e} \boldsymbol{\pi}_{1:n} & p_{22} \boldsymbol{e} \boldsymbol{\pi}_{2:n} & \cdots & p_{2n} \boldsymbol{e} \boldsymbol{\pi}_{n:n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} \boldsymbol{e} \boldsymbol{\pi}_{1:n} & p_{n2} \boldsymbol{e} \boldsymbol{\pi}_{2:n} & \cdots & p_{nn} \boldsymbol{e} \boldsymbol{\pi}_{n:n} \end{pmatrix}. \end{aligned}$$

This and the doubly stochastic nature of $\boldsymbol{P}$ imply that

$$\boldsymbol{\alpha}^{(n)} (-\boldsymbol{S}^{(n)})^{-1} \boldsymbol{D}_{\boldsymbol{P}} = \boldsymbol{\pi}^{(n)}. \qquad (4.4.5)$$

Similarly, we have that

$$\boldsymbol{\pi}^{(n)} (-\boldsymbol{T}^{(n)})^{-1} \boldsymbol{E}_{\boldsymbol{Q}} = \boldsymbol{\alpha}^{(n)}. \qquad (4.4.6)$$

Using (4.4.5) and (4.4.6) alternatingly in (4.4.4), and noting that $\boldsymbol{D_Q e} = \boldsymbol{s}^{(n)}$, we get that

$$\begin{aligned}&\mathbb{P}(S_k \in \mathrm{d}s_k, U_k \in \mathrm{d}u_k, S_{k+1} \in \mathrm{d}s_{k+1})\\&\quad = \boldsymbol{\alpha}^{(n)} e^{\boldsymbol{S}^{(n)} s_k} \boldsymbol{D_P} e^{\boldsymbol{T}^{(n)} u_k} \boldsymbol{E_Q} e^{\boldsymbol{S}^{(n)} s_{k+1}} \boldsymbol{s}^{(n)} \mathrm{d}s_k \mathrm{d}u_k \mathrm{d}s_{k+1}. \qquad (4.4.7)\end{aligned}$$

Integrating (4.4.7) over s_{k+1} implies that

$$\begin{aligned}\mathbb{P}(S_k \in \mathrm{d}s_k, U_k \in \mathrm{d}u_k) &= \boldsymbol{\alpha}^{(n)} e^{\boldsymbol{S}^{(n)} s_k} \boldsymbol{D_P} e^{\boldsymbol{T}^{(n)} u_k} \boldsymbol{E_Q} (-\boldsymbol{S}^{(n)})^{-1} \boldsymbol{s}^{(n)} \mathrm{d}s_k \mathrm{d}u_k\\&= \boldsymbol{\alpha}^{(n)} e^{\boldsymbol{S}^{(n)} s_k} \boldsymbol{D_P} e^{\boldsymbol{T}^{(n)} u_k} \boldsymbol{E_Q} \boldsymbol{e} \mathrm{d}s_k \mathrm{d}u_k\\&= \boldsymbol{\alpha}^{(n)} e^{\boldsymbol{S}^{(n)} s_k} \boldsymbol{D_P} e^{\boldsymbol{T}^{(n)} u_k} \boldsymbol{t}^{(n)} \mathrm{d}s_k \mathrm{d}u_k,\end{aligned}$$

so that (S_k, U_k) follows the bivariate Baker–type distribution

$$H_{\boldsymbol{P}}(x, y) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} F_{i:n}(x) G_{j:m}(y).$$

Integrating (4.4.7) over s_k implies that

$$\begin{aligned}&\mathbb{P}(U_k \in \mathrm{d}u_k, S_{k+1} \in \mathrm{d}s_{k+1})\\&\quad = \boldsymbol{\alpha}^{(n)} (-\boldsymbol{S}^{(n)})^{-1} \boldsymbol{D_P} e^{\boldsymbol{T}^{(n)} u_k} \boldsymbol{E_Q} e^{\boldsymbol{S}^{(n)} s_{k+1}} \boldsymbol{s}^{(n)} \mathrm{d}u_k \mathrm{d}s_{k+1}\\&\quad = \boldsymbol{\pi}^{(n)} e^{\boldsymbol{T}^{(n)} u_k} \boldsymbol{E_Q} e^{\boldsymbol{S}^{(n)} s_{k+1}} \boldsymbol{s}^{(n)} \mathrm{d}u_k \mathrm{d}s_{k+1},\end{aligned}$$

so that (U_k, S_{k+1}) follows the bivariate Baker–type distribution

$$H_{\boldsymbol{Q}}(x, y) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{m} q_{ij} G_{i:n}(x) F_{j:m}(y).$$

Remark 16 *Notice that we are also able to quantify the dependence between S_i and S_{i+1} from (4.4.7). Indeed, if we integrate (4.4.7) over u_k we get*

$$\begin{aligned}&\mathbb{P}(S_k \in \mathrm{d}s_k, S_{k+1} \in \mathrm{d}s_{k+1})\\&\quad = \boldsymbol{\alpha}^{(n)} e^{\boldsymbol{S}^{(n)} s_k} \boldsymbol{D_P} (-\boldsymbol{T}^{(n)})^{-1} \boldsymbol{E_Q} e^{\boldsymbol{S}^{(n)} s_{k+1}} \boldsymbol{s}^{(n)} \mathrm{d}s_k \mathrm{d}s_{k+1}\\&\quad = \boldsymbol{\alpha}^{(n)} e^{\boldsymbol{S}^{(n)} s_k} \boldsymbol{D}^{*} e^{\boldsymbol{S}^{(n)} s_{k+1}} \boldsymbol{s}^{(n)} \mathrm{d}s_k \mathrm{d}s_{k+1},\end{aligned}$$

where

$$\begin{aligned}\boldsymbol{D}^* &= \boldsymbol{D}_{\boldsymbol{P}}(-\boldsymbol{T}^{(n)})^{-1}\boldsymbol{E}_{\boldsymbol{Q}} \\ &= \begin{pmatrix} p_{11}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{1:n} & p_{12}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{1n}\boldsymbol{s}_{1:n}\boldsymbol{\pi}_{n:n} \\ p_{21}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{1:n} & p_{22}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{2n}\boldsymbol{s}_{2:n}\boldsymbol{\pi}_{n:n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{1:n} & p_{n2}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{2:n} & \cdots & p_{nn}\boldsymbol{s}_{n:n}\boldsymbol{\pi}_{n:n} \end{pmatrix} \\ &\quad \times \begin{pmatrix} q_{11}\boldsymbol{e}\boldsymbol{\alpha}_{1:n} & q_{12}\boldsymbol{e}\boldsymbol{\alpha}_{2:n} & \cdots & q_{1n}\boldsymbol{e}\boldsymbol{\alpha}_{n:n} \\ q_{21}\boldsymbol{e}\boldsymbol{\alpha}_{1:n} & q_{22}\boldsymbol{e}\boldsymbol{\alpha}_{2:n} & \cdots & q_{2n}\boldsymbol{e}\boldsymbol{\alpha}_{n:n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1}\boldsymbol{e}\boldsymbol{\alpha}_{1:n} & q_{n2}\boldsymbol{e}\boldsymbol{\alpha}_{2:n} & \cdots & q_{nn}\boldsymbol{e}\boldsymbol{\alpha}_{n:n} \end{pmatrix} \\ &= \begin{pmatrix} p_{11}^*\boldsymbol{s}_{1:n}\boldsymbol{\alpha}_{1:n} & p_{12}^*\boldsymbol{s}_{1:n}\boldsymbol{\alpha}_{2:n} & \cdots & p_{1n}^*\boldsymbol{s}_{1:n}\boldsymbol{\alpha}_{n:n} \\ p_{21}^*\boldsymbol{s}_{2:n}\boldsymbol{\alpha}_{1:n} & p_{22}^*\boldsymbol{s}_{2:n}^*\boldsymbol{\alpha}_{2:n} & \cdots & p_{2n}^*\boldsymbol{s}_{2:n}\boldsymbol{\alpha}_{n:n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1}^*\boldsymbol{s}_{n:n}\boldsymbol{\alpha}_{1:n} & p_{n2}^*\boldsymbol{s}_{n:n}\boldsymbol{\alpha}_{2:n} & \cdots & p_{nn}^*\boldsymbol{s}_{n:n}\boldsymbol{\alpha}_{n:n} \end{pmatrix}, \end{aligned}$$

and $\{p_{ij}^*\} =: \boldsymbol{P}^* := \boldsymbol{PQ}$. *This means that* (S_i, S_{i+1}) *follows the bivariate Baker–type distribution*

$$H_{\boldsymbol{P}^*}(x,y) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^* F_{i:n}(x)F_{j:m}(y).$$

Similarly, it can be verified that (U_i, U_{i+1}) *follows the bivariate Baker–type distribution*

$$H_{\boldsymbol{Q}^*}(x,y) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} q_{ij}^* G_{i:n}(x)G_{j:m}(y),$$

where $\{q_{ij}^*\} = \boldsymbol{Q}^* = \boldsymbol{QP}$.

4.4.3 Sequential

Let $\boldsymbol{P} = \{p_{ij}\}$ and $\boldsymbol{Q} = \{q_{ij}\}$ be $n \times n$ doubly stochastic matrices and let

$$\boldsymbol{\alpha}^{(n)} = n^{-1}(\boldsymbol{\alpha}_{1:n}, \boldsymbol{\alpha}_{2:n}, \ldots, \boldsymbol{\alpha}_{n:n}), \quad \boldsymbol{S}^{(n)} = \begin{pmatrix} \boldsymbol{S}_{1:n} & 0 & \cdots & 0 \\ \boldsymbol{0} & \boldsymbol{S}_{2:n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{S}_{n:n} \end{pmatrix},$$

$$\boldsymbol{\pi}^{(n)} = n^{-1}(\boldsymbol{\pi}_{1:n}, \boldsymbol{\pi}_{2:n}, \ldots, \boldsymbol{\pi}_{n:n}), \quad \boldsymbol{T}^{(n)} = \begin{pmatrix} \boldsymbol{T}_{1:n} & 0 & \cdots & 0 \\ \boldsymbol{0} & \boldsymbol{T}_{2:n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{T}_{n:n} \end{pmatrix},$$

$$\boldsymbol{D_P} = \begin{pmatrix} p_{11}\boldsymbol{s}_{1:n}\boldsymbol{\alpha}_{1:n} & p_{12}\boldsymbol{s}_{1:n}\boldsymbol{\alpha}_{2:n} & \cdots & p_{1n}\boldsymbol{s}_{1:n}\boldsymbol{\alpha}_{n:n} \\ p_{21}\boldsymbol{s}_{2:n}\boldsymbol{\alpha}_{1:n} & p_{22}\boldsymbol{s}_{2:n}\boldsymbol{\alpha}_{2:n} & \cdots & p_{2n}\boldsymbol{s}_{2:n}\boldsymbol{\alpha}_{n:n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1}\boldsymbol{s}_{n:n}\boldsymbol{\alpha}_{1:n} & p_{n2}\boldsymbol{s}_{n:n}\boldsymbol{\alpha}_{2:n} & \cdots & p_{nn}\boldsymbol{s}_{n:n}\boldsymbol{\alpha}_{n:n} \end{pmatrix}, \quad \text{and}$$

$$\boldsymbol{E_Q} = \begin{pmatrix} q_{11}\boldsymbol{t}_{1:n}\boldsymbol{\pi}_{1:n} & q_{12}\boldsymbol{t}_{1:n}\boldsymbol{\pi}_{2:n} & \cdots & q_{1n}\boldsymbol{t}_{1:n}\boldsymbol{\pi}_{n:n} \\ q_{21}\boldsymbol{t}_{2:n}\boldsymbol{\pi}_{1:n} & q_{22}\boldsymbol{t}_{2:n}\boldsymbol{\pi}_{2:n} & \cdots & q_{2n}\boldsymbol{t}_{2:n}\boldsymbol{\pi}_{n:n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1}\boldsymbol{t}_{n:n}\boldsymbol{\pi}_{1:n} & q_{n2}\boldsymbol{t}_{n:n}\boldsymbol{\pi}_{2:n} & \cdots & q_{nn}\boldsymbol{t}_{n:n}\boldsymbol{\pi}_{n:n} \end{pmatrix}.$$

Let $\mathfrak{E}_{S^{(n)}}$ and $\mathfrak{E}_{T^{(n)}}$ be the phase–spaces associated to $\boldsymbol{S}^{(n)}$ and $\boldsymbol{T}^{(n)}$, respectively. Let $\{R_t\}_{t\geq 0}$ be the fluid flow risk process with associated fluid flow process $\{(V_t, J_t)\}_{t\geq 0}$, where $\{J_t\}_{t\geq 0}$ has state–space $\mathfrak{E}_+ \cup \mathfrak{E}_-$ with $\mathfrak{E}_+ = \mathfrak{E}_{S^{(n)}} \times \mathfrak{E}_{T^{(n)}}$ and $\mathfrak{E}_- = \mathfrak{E}_{S^{(n)}} \times \mathfrak{E}_{T^{(n)}}$ (with $\mathfrak{E}_+$ and $\mathfrak{E}_+$ regarded as disjoint sets), $J_0 \sim (\boldsymbol{\alpha}^{(n)} \otimes \boldsymbol{\pi}^{(n)}, \boldsymbol{0})$ and has an intensity matrix given by

$$\begin{pmatrix} \boldsymbol{S}^{(n)} \otimes \boldsymbol{I} & \boldsymbol{D_P} \otimes \boldsymbol{I} \\ \boldsymbol{I} \otimes \boldsymbol{E_Q} & \boldsymbol{I} \otimes \boldsymbol{T}^{(n)} \end{pmatrix}, \tag{4.4.8}$$

where $\otimes$ denotes the Kronecker product (see Appendix A). Let the reward rates $\{r_i\}_{i\in\mathfrak{E}_+\cup\mathfrak{E}_-}$ and Brownian noise rates $\{\sigma_i\}_{i\in\mathfrak{E}_+\cup\mathfrak{E}_-}$ be given by

$$\begin{aligned} r_i = p \quad &\text{and} \quad \sigma_i = 0, \quad i \in \mathfrak{E}_+, \\ r_i = -1 \quad &\text{and} \quad \sigma_i = 0, \quad i \in \mathfrak{E}_-. \end{aligned}$$

We claim that $\{R_t\}_{t\geq 0}$ corresponds to a Sparre–Andersen process with sequential dependence. Indeed, the structure of (4.4.8) implies that

$$\begin{aligned} &\mathbb{P}(S_1 \in \mathrm{d}s_1, U_1 \in \mathrm{d}u_1, S_2 \in \mathrm{d}s_2, U_2 \in \mathrm{d}u_2, \ldots, S_k \in \mathrm{d}s_k, U_k \in \mathrm{d}u_k) \\ &= (\boldsymbol{\alpha}^{(n)} \otimes \boldsymbol{\pi}^{(n)}) \left(\prod_{i=1}^{k} e^{\boldsymbol{S}^{(n)}\otimes \boldsymbol{I} s_i} (\boldsymbol{D_P} \otimes \boldsymbol{I}) e^{\boldsymbol{I}\otimes \boldsymbol{T}^{(n)} u_i} (\boldsymbol{I} \otimes \boldsymbol{E_Q}) \right) \boldsymbol{e} \\ &\qquad \times \mathrm{d}s_1 \mathrm{d}u_1 \mathrm{d}s_2 \mathrm{d}u_2 \ldots \mathrm{d}s_k \mathrm{d}u_k \\ &= (\boldsymbol{\alpha}^{(n)} \otimes \boldsymbol{\pi}^{(n)}) \left(\prod_{i=1}^{k} (e^{\boldsymbol{S}^{(n)} s_i} \otimes \boldsymbol{I})(\boldsymbol{D_P} \otimes \boldsymbol{I})(\boldsymbol{I} \otimes e^{\boldsymbol{T}^{(n)} u_i})(\boldsymbol{I} \otimes \boldsymbol{E_Q}) \right) (\boldsymbol{e} \otimes \boldsymbol{e}) \\ &\qquad \times \mathrm{d}s_1 \mathrm{d}u_1 \mathrm{d}s_2 \mathrm{d}u_2 \ldots \mathrm{d}s_k \mathrm{d}u_k \\ &= \left[\boldsymbol{\alpha}^{(n)} \left(\prod_{i=1}^{k} e^{\boldsymbol{S}^{(n)} s_i} \boldsymbol{D_P} \right) \boldsymbol{e} \mathrm{d}s_1 \ldots \mathrm{d}s_k \right] \left[\boldsymbol{\pi}^{(n)} \left(\prod_{i=1}^{k} e^{\boldsymbol{T}^{(n)} u_i} \boldsymbol{E_Q} \right) \boldsymbol{e} \mathrm{d}u_1 \ldots \mathrm{d}u_k \right], \end{aligned}$$

where the second equality follows from Proposition A.2 and the last one from Proposition A.1. This implies that $\{S_i\}_{i\geq 1}$ is independent from $\{U_i\}_{i\geq 1}$, and that

$$\mathbb{P}(S_1 \in \mathrm{d}s_1, S_2 \in \mathrm{d}s_2, \ldots, S_k \in \mathrm{d}s_k) = \boldsymbol{\alpha}^{(n)} \left(\prod_{i=1}^{k} e^{\boldsymbol{S}^{(n)} s_i} \boldsymbol{D_P} \right) \boldsymbol{e} \mathrm{d}s_1 \ldots \mathrm{d}s_k, \tag{4.4.9}$$

$$\mathbb{P}(U_1 \in \mathrm{d}u_1, U_2 \in \mathrm{d}u_2, \ldots, U_k \in \mathrm{d}u_k) = \boldsymbol{\pi}^{(n)} \left(\prod_{i=1}^{k} e^{\boldsymbol{T}^{(n)} u_i} \boldsymbol{E_Q} \right) \boldsymbol{e} \mathrm{d}u_1 \ldots \mathrm{d}u_k. \tag{4.4.10}$$

Integrating (4.4.9) with respect to $s_1, \ldots, s_{k-2}$, integrating (4.4.10) with respect to $u_1, \ldots, u_{k-2}$, and using that

$$\boldsymbol{\alpha}^{(n)}(-\boldsymbol{S}^{(n)})^{-1}\boldsymbol{D_P} = \boldsymbol{\alpha}^{(n)}, \quad \boldsymbol{\pi}^{(n)}(-\boldsymbol{T}^{(n)})^{-1}\boldsymbol{E_Q} = \boldsymbol{\pi}^{(n)},$$

$$\boldsymbol{D_P e} = \boldsymbol{s}^{(n)}, \quad \text{and} \quad \boldsymbol{E_Q e} = \boldsymbol{t}^{(n)},$$

we get that

$$\begin{aligned}
\mathbb{P}(S_{k-1} \in \mathrm{d}s_{k-1}, S_k \in \mathrm{d}s_k) &= \boldsymbol{\alpha}^{(n)} e^{\boldsymbol{S}^{(n)} s_i} \boldsymbol{D_P} e^{\boldsymbol{S}^{(n)} s_i} \boldsymbol{s}^{(n)} \mathrm{d}s_{k-1} \mathrm{d}s_k, \\
\mathbb{P}(U_{k-1} \in \mathrm{d}u_{k-1}, U_k \in \mathrm{d}u_k) &= \boldsymbol{\pi}^{(n)} e^{\boldsymbol{T}^{(n)} u_i} \boldsymbol{E_Q} e^{\boldsymbol{T}^{(n)} u_i} \boldsymbol{t}^{(n)} \mathrm{d}u_{k-1} \mathrm{d}u_k.
\end{aligned}$$

This implies that (S_{k-1}, S_k) follows the bivariate Baker–type distribution

$$H_{\boldsymbol{P}}(x, y) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} F_{i:n}(x) F_{j:m}(y),$$

and that (U_{k-1}, U_k) follows the bivariate Baker–type distribution

$$H_{\boldsymbol{Q}}(x, y) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{m} q_{ij} G_{i:n}(x) G_{j:m}(y).$$

We refer to Bladt et al. (2018) for an exhaustive numerical assessment on the effects of dependency for the infinite–horizon probability of ruin of dependent Sparre–Andersen models.

4.5 Conclusions and remarks

In this chapter we constructed a class of bivariate phase–type distributions with arbitrary coefficient of correlation, and proposed an application of it in risk theory to model processes with dependencies. Our constructions, both the bivariate phase–type distributions and models with dependencies, compare to existing results and models in the literature as follows.

- The bivariate phase–type distribution of the Baker–type we construct in Theorem 4.9 has an explicit density given by (4.3.2). In comparison, there is no explicit density available for a general MPH* distribution: in Kulkarni (1989) the density is characterised by a system of differential partial equations, and in Breuer (2016) a semi–explicit form is deduced.
- A method to fit a given bivariate distribution with a MPH* distribution is proposed in Breuer (2016), as a generalization of the EM algorithm in Asmussen et al. (1996). The method we propose in Theorem 4.9 is more of a straightforward calibration rather than a statistical fitting.
- The Pearson's correlation coefficient is a naive way to measure dependence. This means that in certain cases, choosing $\boldsymbol{P}$ on the form (4.2.8) might provide a poor fit of the goal bivariate distribution. An alternative approach would be to choose $\boldsymbol{P}$ in such a way that $H_{\boldsymbol{P}}$ is indeed close to the goal bivariate distribution, not only to its correlation coefficient. This can be done by considering the copula associated to the Baker–type bivariate distribution (4.2.3), which was defined in Sancetta and Satchell (2004) as Bernstein copula. In Sancetta and Satchell (2004) it is proved that the class of Bernstein copulas can be used to approximate arbitrarily well (with respect to some metric) any copula; a method to construct Bernstein copula approximations based on the EM algorithm is proposed in Dou et al. (2016). Reverting back from Bernstein copula to its multivariate analogue, this means that we can approximate any dependence arbitrarily "well" with a Baker–type bivariate distribution (4.2.3) by choosing an adequate n (the amount of order statistics used for F and G) and $\boldsymbol{P}$. Considering that the class of phase–type distributions can approximate arbitrarily well any distribution with support on $\mathbb{R}_+$ (see Bladt and Nielsen (2017)), we conclude that the class of bivariate distributions constructed in Theorem 4.9 are dense (with respect to some metric) within the set of bivariate distributions with support on $\mathbb{R}_+^2$.
- The dimension of the underlying Markov jump process in (4.3.1) gets large quickly, even for moderate values of n. More precisely, if p_1 is the dimension of the phase–type representation $\text{PH}(\boldsymbol{\alpha}, \boldsymbol{S})$ corresponding to F, and p_2 is the dimension of the phase–type representation $\text{PH}(\boldsymbol{\pi}, \boldsymbol{T})$

corresponding to G, then the dimension d of the underlying Markov jump process of (4.3.1) is given by

$$d = \sum_{i=1}^{n} i(p_1^i + p_2^i).$$

For instance, if $p_1 = p_2 = 3$, then $d = 204,852,3282$ for $n = 3, 4, 5$. This can cause computational issues, which limits the extent to which Theorem 4.9 can actually be used for real–life applications.

- The setting of dependent Sparre–Andersen processes (or its equivalent queueing model) has been considered before in several scenarios. For instance, the study of dependence by pairs goes back at least to Conolly (1968) for the case of a $M/M/1$ queue. Later on, dependence by pairs was pursued for more general models in Borst et al. (1992) for the $M/G/1$ queue, and in Müller (2000) for the $G/G/1$ queue. In the risk theory literature, dependence by pairs was studied by Albrecher and Boxma (2004) and in Boudreault et al. (2006) for the Cramér—Lundberg process, later to be extended to the Sparre-Andersen process in Albrecher et al. (2014). Matrix–analytic methods have been used before to study Sparre–Andersen processes with dependence by pairs, for instance, in Badescu et al. (2009), Badila et al. (2014), Badila et al. (2015) and Avram et al. (2016): in some sense, our model in Subsection 4.4.1 overlaps with these works. To the best of our knowledge, the models of Subsections 4.4.2 and 4.4.3, that is, the Sparre–Andersen processes with alternating and sequential dependence, have not been explored in the literature.

Chapter 5

Parisian ruin for fluid flow risk processes

5.1 Introduction

A risk process is declared ruined within an infinite–horizon if it eventually downcrosses level 0; this is not the case in a Parisian nor cumulative Parisian setting. Parisian ruin (introduced by Dassios and Wu (2008)) happens if a risk process, upon becoming negative, does not recover within some prescribed random time which is restarted at the beginning of each subexcursion below 0. On the other hand, cumulative Parisian ruin happens if the total occupation time of the risk process below 0 is greater than some prescribed random time. In this chapter we compute the probability that a fluid flow risk process gets ruined in a Parisian or cumulative Parisian way under the assumption that the clocks are phase–type–distributed. Our analysis is done for two cases: for fluid flow risk processes that do not have Brownian components, and for fluid flow risk processes that have Brownian components. We propose an erlangization method similar to the one of Asmussen et al. (2002) to approximate the probability of Parisian and cumulative Parisian ruin with deterministic clocks. We finalise with a discussion and comparison of our results with previous work done in the area.

The main original results stemming from this chapter are Corollary 5.7, Theorem 5.9, Corollary 5.18, Theorem 5.20, Theorem 5.21, Theorem 5.22 and Theorem 5.23.

5.2 Fluid flow risk process without Brownian noise

In this section we compute the probability of ruin in an infinite–horizon, Parisian and cumulative Parisian for a fluid flow risk process with no Brownian components.

5.2.1 Infinite–horizon probability of ruin

Let $\{R_t\}_{t\geq 0}$ be a fluid flow risk process and let $\{(V_t, J_t)\}_{t\geq 0}$ be its underlying fluid flow process with $\mathfrak{E}_\sigma = \emptyset$. This means that the intensity matrix of $\{J_t\}_{t\geq 0}$ can be written on the form

$$\mathbf{\Lambda} = \begin{pmatrix} \mathbf{\Lambda}_{++} & \mathbf{\Lambda}_{+-} & \mathbf{\Lambda}_{+0} \\ \mathbf{\Lambda}_{-+} & \mathbf{\Lambda}_{--} & \mathbf{\Lambda}_{-0} \\ \mathbf{\Lambda}_{0+} & \mathbf{\Lambda}_{0-} & \mathbf{\Lambda}_{00} \end{pmatrix},$$

the reward rates $\{r_i\}_{i\in\mathfrak{E}_+\cup\mathfrak{E}_-\cup\mathfrak{E}_0}$ are such that

$$\begin{aligned} r_i &> 0 \quad \text{for all} \quad i \in \mathfrak{E}_+, \\ r_i &< 0 \quad \text{for all} \quad i \in \mathfrak{E}_-, \\ r_i &= 0 \quad \text{for all} \quad i \in \mathfrak{E}_0, \end{aligned}$$

and the Brownian noise rates $\{\sigma_i\}_{i\in\mathfrak{E}_+\cup\mathfrak{E}_-\cup\mathfrak{E}_0}$ are all 0. Furthermore, assume that $R_t \to +\infty$ as $t \to \infty$, that is, the process $\{R_t\}_{t\geq 0}$ has positive drift.

Let $\mathbf{\Lambda}_c$ correspond to the intensity matrix of the $\mathfrak{E}_0$–censored process $\{J_t\}_{t\geq 0}$, so that

$$\mathbf{\Lambda}_c = \begin{pmatrix} \mathbf{\Lambda}_{++} + \mathbf{\Lambda}_{+0}(-\mathbf{\Lambda}_{00})^{-1}\mathbf{\Lambda}_{0+} & \mathbf{\Lambda}_{+-} + \mathbf{\Lambda}_{+0}(-\mathbf{\Lambda}_{00})^{-1}\mathbf{\Lambda}_{0-} \\ \mathbf{\Lambda}_{-+} + \mathbf{\Lambda}_{-0}(-\mathbf{\Lambda}_{00})^{-1}\mathbf{\Lambda}_{0+} & \mathbf{\Lambda}_{--} + \mathbf{\Lambda}_{-0}(-\mathbf{\Lambda}_{00})^{-1}\mathbf{\Lambda}_{0-} \end{pmatrix} =: \begin{pmatrix} \mathbf{\Lambda}^c_{++} & \mathbf{\Lambda}^c_{+-} \\ \mathbf{\Lambda}^c_{-+} & \mathbf{\Lambda}^c_{--} \end{pmatrix}.$$

By Theorem 2.31 and Theorem 2.33, we know that if $\tau_0 = \inf\{t > 0 : V_t < 0\}$ and $m_0 = J_{\tau_0}$,

$$\mathbb{P}(m_0 = j, \tau_0 < \infty \mid J_0 = i, V_0 = u) = \left(\boldsymbol{e}_i' \boldsymbol{\beta}^c_{+-} e^{\boldsymbol{D}^c u}\right)_j, \quad i \in \mathfrak{E}_+, j \in \mathfrak{E}_-, u \geq 0,$$

where $\boldsymbol{\beta}^c_{+-}$ and $\boldsymbol{D}^c$ satisfy the equations

$$\boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}^c_{++}\boldsymbol{\beta}^c_{+-} + \boldsymbol{\beta}^c_{+-}\boldsymbol{\Lambda}^c_{--} + \boldsymbol{\beta}^c_{+-}\boldsymbol{\Lambda}^c_{-+}\boldsymbol{\beta}^c_{+-} + \boldsymbol{\Delta}_{r_+}^{-1}\boldsymbol{\Lambda}^c_{+-} = \mathbf{0}, \quad \text{and} \tag{5.2.1}$$

$$\boldsymbol{D}^c = \boldsymbol{\Lambda}^c_{--} + \boldsymbol{\Lambda}^c_{-+}\boldsymbol{\beta}^c_{+-}, \tag{5.2.2}$$

and which exact value can be computed via the iterative algorithm of Theorem 2.32 or via any algorithm from Bean et al. (2005). Since $\{R_t\}_{t\geq 0}$ and $\{V_t\}_{t\geq 0}$ share the same first passage probabilities, we have the following.

Theorem 5.1 *The infinite–horizon probability of ruin defined by*

$$\psi_i(u) := \mathbb{P}(\inf_t R_t < 0 \mid R_0 = 0, J_0 = i)$$

is given by

$$\psi_i(u) = \boldsymbol{e}_i'\boldsymbol{\beta}^c_{+-}e^{\boldsymbol{D}^c u}\boldsymbol{e}, \quad i \in \mathfrak{E}_+, u \geq 0.$$

5.2.2 Parisian and cumulative Parisian ruin

Define $\{d_n\}_{n\geq 1}$ and $\{u_n\}_{n\geq 1}$ as the (possibly finite) set of sequential points in time at which the process $\{V_t\}_{t\geq 0}$ downcrosses and upcrosses 0, respectively. More precisely, we let $u_0 = 0$ and for $n \geq 1$,

$$\begin{aligned} d_n &= \inf\{t \geq u_{n-1} : V_t < 0\} \\ u_n &= \inf\{t > d_n : V_t \geq 0\}, \end{aligned}$$

so that $\{[u_{n-1}, d_n)\}_{n\geq 1}$ corresponds to the sequence of excursions of $\{V_t\}_{t\geq 0}$ above zero, and $\{[d_n, u_n)\}_{n\geq 1}$ corresponds to the sequence of excursions of $\{V_t\}_{t\geq 0}$ below zero. For all $n \geq 1$, let

$$h_n = \int_{d_n}^{u_n} \mathbb{1}_{J_s \in E_+} \mathrm{d}s$$

($h_n = 0$ for $d_n = +\infty$) denote the total amount of actual time, in the sense of the equivalent process $\{R_t\}_{t\geq 0}$, accumulated in the n-th subexcursion below 0. See Figure 5.1 for a visual description of $\{h_n\}_{n\geq 1}$.

Definition 5.2 (Parisian ruin) We say that the fluid flow risk process $\{R_t\}_{t\geq 0}$ (with $\mathfrak{E}_\sigma = \emptyset$ and underlying fluid flow process $\{(V_t, J_t)\}_{t\geq 0}$) gets ruined in the **Parisian** way with F–distributed clocks if there exists some $n \geq 1$ such that $L_n < h_n$, where $\{L_i\}_{i\geq 1}$ is an independent sequence of nonnegative i.i.d.r.v.'s with common distribution F. For any $i \in \mathfrak{E}_+$ and $u \geq 0$, we define the probability of Parisian ruin $\psi_i^F(u)$ by

$$\psi_i^F(u) = \mathbb{P}\left(\cup_{n\geq 1}\{L_n < h_n\} \mid V_0 = u, J_0 = i\right).$$

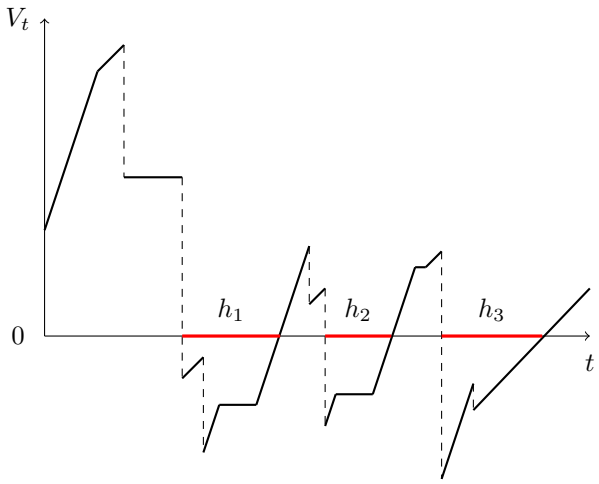


Figure 5.1: A fluid flow risk process $\{R_t\}_{t\geq 0}$ with $\mathfrak{E}_\sigma = \emptyset$. The intervals in red correspond to the succesive subexcursions of $\{R_t\}_{t\geq 0}$ below 0, with lengths h_i, $i \geq 1$.

Definition 5.3 (Cumulative Parisian ruin) We say that the fluid flow risk process $\{R_t\}_{t\geq 0}$ (with $\mathfrak{E}_\sigma = \emptyset$ and underlying fluid flow process $\{(V_t, J_t)\}_{t\geq 0}$) gets ruined in the **cumulative Parisian** way with an F–distributed clock if $L \leq \sum_{n=1}^{\infty} h_n$, where L is an independent r.v. with distribution F. For any $i \in \mathfrak{E}_+$ and $u \geq 0$, we define the probability of cumulative Parisian ruin $\phi_i^F(u)$ by

$$\phi_i^F(u) = \mathbb{P}\left(L \leq \sum_{n=1}^{\infty} h_n \mid V_0 = u, J_0 = i \right).$$

In the Parisian setting, a clock L_n is generated if and when $\{R_t\}_{t\geq 0}$ downcrosses 0 for the n-th time. If $\{R_t\}_{t\geq 0}$ can manage to get back into positive before h_n exceeds L_n, then Parisian ruin has been avoided in this sub–excursion below 0. The process $\{R_t\}_{t\geq 0}$ may thus get ruined in the infinite–horizon sense, possibly become negative several times, but still survive in the Parisian setting. In the cumulative Parisian setting, $\{R_t\}_{t\geq 0}$ is not declared ruined as long as $\int_0^\infty \mathbb{1}\{R_s \leq 0\}\mathrm{d}s < L$, even if it downcrosses 0 several times.

We study the case of Parisian ruin first. Consider the first downcrossing of 0 by $\{V_t\}_{t\geq 0}$. By the strong Markov property of $\{(V_t, J_t)\}_{t\geq 0}$, it is enough to study the scenario in which $V_0 = 0$ and $J_0 = i \in \mathfrak{E}_-$. Under these assumptions we

define

$$\tau = \inf\{t > 0 : V_t = 0, J_t \notin \mathfrak{E}_-\} \text{ and } h = \int_0^\tau \mathbb{1}\{V_s < 0, J_s \notin \mathfrak{E}_-\}\mathrm{d}s. \quad (5.2.3)$$

For $i \in \mathfrak{E}_-$ and $j \in \mathfrak{E}_+$, we let

$$\bar{\psi}_{ij}(F) = \mathbb{P}(L > h, J_\tau = j \mid V_0 = 0, J_0 = i), \quad L \sim F, L \perp \{(V_t, J_t)\}_{t\geq 0},$$

and let $\bar{\boldsymbol{\Psi}}(F) = \{\bar{\psi}_{ij}(F)\}_{i\in\mathfrak{E}_-, j\in\mathfrak{E}_+}$ denote the corresponding matrix. Then have the following result.

Proposition 5.4 *Let $u \geq 0$, $i \in \mathfrak{E}_+$ and $n \geq 1$. Then,*

$$\begin{aligned}\mathbb{P}&\left(\bigcap_{j<n}\{L_j > h_j\}, L_n < h_n \mid V_0 = u, J_0 = i\right)\\ &= \boldsymbol{e}_i'\boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u}\left(\bar{\boldsymbol{\Psi}}(F)\boldsymbol{\beta}_{+-}^c\right)^{n-1}(\boldsymbol{e} - \bar{\boldsymbol{\Psi}}(F)\boldsymbol{e}). \end{aligned} \quad (5.2.4)$$

Proof. For $n = 1$, the event $\{\bigcap_{j<n}\{L_j > h_j\}, L_n < h_n\} = \{L_1 < h_1\}$ implies that $\{R_t\}_{t\geq 0}$ downcrossed level 0 and it was not able to upcross it before L_1 rang. Conditioning on the state of the first downcrossing of level 0, we get that

$$\begin{aligned}\mathbb{P}\left(L_1 < h_1 \mid V_0 = u, J_0 = i\right) &= \sum_{j\in\mathfrak{E}_-}(\boldsymbol{e}_i'\boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u})_j(1 - \boldsymbol{e}_j'\bar{\boldsymbol{\Psi}}(F)\boldsymbol{e})\\ &= \boldsymbol{e}_i'\boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u}(\boldsymbol{e} - \bar{\boldsymbol{\Psi}}(F)\boldsymbol{e})\end{aligned}$$

For $n \geq 2$, the event $\{\bigcap_{j<n}\{L_j > h_j\}, L_n < h_n\}$ implies that $\{R_t\}_{t\geq 0}$:

1. Downcrossed 0, which happens according to the probability vector $\boldsymbol{e}_i'\boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u}$;
2. For $j \in \{1, \ldots, n-1\}$,
   (a) Uprossed 0 before the clock L_j rang, which happens according to the probability matrix $\bar{\boldsymbol{\Psi}}(F)$;
   (b) Downcrossed level 0, which happens according to the probability matrix $\boldsymbol{\beta}_{+-}^c$;
3. Did not upcross level 0 before L_n rang, which happens according to the probability vector $\boldsymbol{e} - \bar{\boldsymbol{\Psi}}(F)\boldsymbol{e}$.

Thus, (5.2.4) follows. □

Theorem 5.5 *For any $i \in \mathfrak{E}_+$ and $u \geq 0$, the probability of Parisian ruin $\psi_i^F(u)$ is given by*

$$\psi_i^F(u) = \boldsymbol{e}_i' \boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u} \left(\boldsymbol{I} - \bar{\boldsymbol{\Psi}}(F)\boldsymbol{\beta}_{+-}^c\right)^{-1} (\boldsymbol{e} - \bar{\boldsymbol{\Psi}}(F)\boldsymbol{e}).$$

Proof. Simply notice that

$$\begin{aligned}
\psi_i^F(u) &= \sum_{n=1}^{\infty} \mathbb{P}\left(\bigcap_{j<n} \{L_j > h_j\}, L_n < h_n \mid V_0 = u, J_0 = i\right) \\
&= \sum_{n=1}^{\infty} \boldsymbol{e}_i' \boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u} \left(\bar{\boldsymbol{\Psi}}(F)\boldsymbol{\beta}_{+-}^c\right)^{n-1} (\boldsymbol{e} - \bar{\boldsymbol{\Psi}}(F)\boldsymbol{e}) \\
&= \boldsymbol{e}_i' \boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u} \left(\sum_{n=0}^{\infty} \left(\bar{\boldsymbol{\Psi}}(F)\boldsymbol{\beta}_{+-}^c\right)^n\right) (\boldsymbol{e} - \bar{\boldsymbol{\Psi}}(F)\boldsymbol{e}) \\
&= \boldsymbol{e}_i' \boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u} \left(\boldsymbol{I} - \bar{\boldsymbol{\Psi}}(F)\boldsymbol{\beta}_{+-}^c\right)^{-1} (\boldsymbol{e} - \bar{\boldsymbol{\Psi}}(F)\boldsymbol{e}). \qquad (5.2.5)
\end{aligned}$$

The positive drift assumption of $\{V_t\}_{t\geq 0}$ implies that $\left(\boldsymbol{\beta}_{+-}^c \boldsymbol{e}\right)_j < 1$ for all $j \in \mathfrak{E}_+$. This implies that $\bar{\boldsymbol{\Psi}}(F)\boldsymbol{\beta}_{+-}^c$ is a sub–transition matrix, $\bar{\boldsymbol{\Psi}}(F)\boldsymbol{\beta}_{+-}^c - \boldsymbol{I}$ is a sub–intensity matrix, and Theorem 2.8 implies that $(\boldsymbol{I} - \bar{\boldsymbol{\Psi}}(F)\boldsymbol{\beta}_{+-}^c)^{-1}$ indeed exists. □

For F corresponding to a phase–type distribution, we obtain an explicit formula for $\bar{\boldsymbol{\Psi}}(F)$ as follows. Let $L \sim \mathrm{PH}_\ell(\boldsymbol{\kappa}, \boldsymbol{K})$ and denote by $\mathfrak{E}_\ell$ its phase–space. Let $\{R_t^*\}_{t\geq 0}$ be a fluid flow risk process with an underlying fluid flow process $\{(V_t^*, J_t^*)\}_{t\geq 0}$ where $\{J_t^*\}_{t\geq 0}$ has state–space $\mathfrak{E}_+^* \cup \mathfrak{E}_-^* \cup \mathfrak{E}_0^*$ with $\mathfrak{E}_+^* = \mathfrak{E}_+ \times \mathfrak{E}_\ell$, $\mathfrak{E}_-^* = \mathfrak{E}_- \times \mathfrak{E}_\ell$ and $\mathfrak{E}_0^* = \mathfrak{E}_0 \times \mathfrak{E}_\ell$, ordered lexicographically when necessary. The intensity matrix of $\{J_t^*\}_{t\geq 0}$ is given by

$$\boldsymbol{\Lambda}^* = \begin{pmatrix} \boldsymbol{\Lambda}_{++} \oplus \boldsymbol{K} & \boldsymbol{\Lambda}_{+-} \otimes \boldsymbol{I} & \boldsymbol{\Lambda}_{+0} \otimes \boldsymbol{I} \\ \boldsymbol{\Lambda}_{-+} \otimes \boldsymbol{I} & \boldsymbol{\Lambda}_{--} \otimes \boldsymbol{I} & \boldsymbol{\Lambda}_{-0} \otimes \boldsymbol{I} \\ \boldsymbol{\Lambda}_{0+} \otimes \boldsymbol{I} & \boldsymbol{\Lambda}_{0-} \otimes \boldsymbol{I} & \boldsymbol{\Lambda}_{00} \oplus \boldsymbol{K} \end{pmatrix} =: \begin{pmatrix} \boldsymbol{\Lambda}_{++}^* & \boldsymbol{\Lambda}_{+-}^* & \boldsymbol{\Lambda}_{+0}^* \\ \boldsymbol{\Lambda}_{-+}^* & \boldsymbol{\Lambda}_{--}^* & \boldsymbol{\Lambda}_{-0}^* \\ \boldsymbol{\Lambda}_{0+}^* & \boldsymbol{\Lambda}_{0-}^* & \boldsymbol{\Lambda}_{00}^* \end{pmatrix}, \quad (5.2.6)$$

and the reward rates $\{r_{(i,j)}^*\}_{(i,j)\in\mathfrak{E}_+^*\cup\mathfrak{E}_-^*\cup\mathfrak{E}_0^*}$ and Brownian noise rates $\{\sigma_{(i,j)}^*\}_{(i,j)\in\mathfrak{E}_+^*\cup\mathfrak{E}_-^*\cup\mathfrak{E}_0^*}$ are given by

$$r_{(i,j)}^* = r_i \quad \text{and} \quad \sigma_{(i,j)}^* = 0, \quad (i,j) \in \mathfrak{E}_+^* \cup \mathfrak{E}_-^* \cup \mathfrak{E}_0^*.$$

Let

$$J_0^* \sim \boldsymbol{e}_i' \otimes \boldsymbol{\kappa} \qquad (5.2.7)$$

be the initial distribution of $\{J_t^*\}_{t\geq 0}$, for some fixed $i \in \mathfrak{E}_+ \cup \mathfrak{E}_- \cup \mathfrak{E}_0$. The behaviour of $\{(V_t^*, J_t^*)\}_{t\geq 0}$ is explained as follows. Write $J_t^* = (i_t^1, i_t^2)$. By

(5.2.7), we have that $i_0^1 = i$ and $i_0^2 \sim \boldsymbol{\kappa}$. According to (5.2.6), $\{i_t^2\}_{t\geq 0}$ will remain fixed or frozen for all $t \geq 0$ such that $i_t^1 \in \mathfrak{E}_-$. While $i_t^1 \in \mathfrak{E}_+ \cup \mathfrak{E}_0$, $\{i_t^2\}_{t\geq 0}$ will develop according to the sub–intensity matrix $\boldsymbol{K}$, so that $\{i_t^2\}_{t\geq 0}$ is a terminating process. On the other hand, $\{i_t\}_{t\geq 0}$ will simply develop according to the original matrix $\boldsymbol{\Lambda}$ up to the termination of $\{i_t^2\}_{t\geq 0}$. See Figure 5.2 for an example of a realization of the process $\{(V_t^*, J_t^*)\}_{t\geq 0}$. All in all, this means

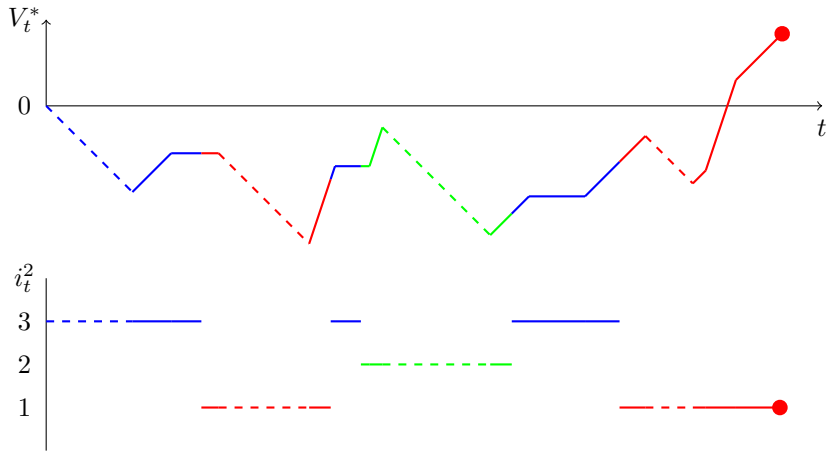


Figure 5.2: An example of the process $\{(V_t^*, J_t^*)\}_{t\geq 0}$ with $V_t = (i_t^1, i_t^2), t \geq 0$. Notice that the process $\{i_t^2\}_{t\geq 0}$ does **not** jump while $i_t^1 \in \mathfrak{E}_-$, that is, at the dashed intervals. The total lifetime of $\{i_t^2\}_{t\geq 0}$ minus the dashed intervals corresponds to the length of the random clock L.

that the fluid flow risk process $\{R_t^*\}_{t\geq 0}$ has the distribution of $\{R_t\}_{t\geq 0}$ killed at time L, in particular, it has the same first passage probabilities.

Now, let $\{R_t^{*c}\}_{t\geq 0}$ be the fluid flow risk process with underlying fluid flow process $\{(V_t^{*c}, J_t^{*c})\}_{t\geq 0}$ which results from censoring the states $\mathfrak{E}_0^*$ in $\{J_t^*\}$. Then, $\{J_t^{*c}\}$ has a state space $\mathfrak{E}_+^* \cup \mathfrak{E}_+^*$ and is driven by the sub–intensity matrix

$$\boldsymbol{\Lambda}^{*c} = \begin{pmatrix} \boldsymbol{\Lambda}_{++}^* + \boldsymbol{\Lambda}_{+0}^*(-\boldsymbol{\Lambda}_{00}^*)^{-1}\boldsymbol{\Lambda}_{0+}^* & \boldsymbol{\Lambda}_{+-}^* + \boldsymbol{\Lambda}_{+0}^*(-\boldsymbol{\Lambda}_{00}^*)^{-1}\boldsymbol{\Lambda}_{0-}^* \\ \boldsymbol{\Lambda}_{-+}^* + \boldsymbol{\Lambda}_{-0}^*(-\boldsymbol{\Lambda}_{00}^*)^{-1}\boldsymbol{\Lambda}_{0+}^* & \boldsymbol{\Lambda}_{--}^* + \boldsymbol{\Lambda}_{-0}^*(-\boldsymbol{\Lambda}_{00}^*)^{-1}\boldsymbol{\Lambda}_{0-}^* \end{pmatrix} =: \begin{pmatrix} \boldsymbol{\Lambda}_{++}^{*c} & \boldsymbol{\Lambda}_{+-}^{*c} \\ \boldsymbol{\Lambda}_{-+}^{*c} & \boldsymbol{\Lambda}_{--}^{*c} \end{pmatrix}. \tag{5.2.8}$$

The process $\{(V_t^{*c}, J_t^{*c})\}_{t\geq 0}$ has reward rates $\{r_{(i,j)}^*\}_{(i,j)\in\mathfrak{E}_+^*\cup\mathfrak{E}_-^*}$ and Brownian noise rates $\{\sigma_{(i,j)}^*\}_{(i,j)\in\mathfrak{E}_+^*\cup\mathfrak{E}_-^*}$ given by

$$r_{(i,j)}^* = r_i \quad \text{and} \quad \sigma_{(i,j)}^* = 0, \quad (i,j) \in \mathfrak{E}_+^* \cup \mathfrak{E}_-^*.$$

Since $\{(V_t^{*c}, J_t^{*c})\}_{t\geq 0}$ is the result of censoring the states $\mathfrak{E}_0^*$ in $\{J_t^*\}_{t\geq 0}$, then the first passage probabilities between $\{R_t^{*c}\}_{t\geq 0}$ and $\{R_t^c\}_{t\geq 0}$ remain the same.

This also means that $\{R_t^{*c}\}_{t\geq 0}$ has the same first passage probabilities as the process $\{R_t\}_{t\geq 0}$ killed at time L.

Let $\boldsymbol{\alpha}_{-+}^{*c}$ be the matrix of upcrossing probabilities of $\{(V_t^{*c}, J_t^{*c})\}_{t\geq 0}$ and let $\boldsymbol{U}^{*c}$ be the intensity matrix of its upward crossing process. More specifically, let $\boldsymbol{\alpha}_{-+}^{*c}$ and $\boldsymbol{U}^{*c}$ be solutions of

$$\boldsymbol{\Lambda}_{--}^{*c}\boldsymbol{\alpha}_{-+}^{*c} + \boldsymbol{\alpha}_{-+}^{*c}\boldsymbol{\Delta}_{r_+^*}^{-1}\boldsymbol{\Lambda}_{++}^{*c} + \boldsymbol{\alpha}_{-+}^{*c}\boldsymbol{\Delta}_{r_+^*}^{-1}\boldsymbol{\Lambda}_{+-}^{*c}\boldsymbol{\alpha}_{-+}^{*c} + \boldsymbol{\Lambda}_{-+}^{*c} = \boldsymbol{0}, \quad \text{and} \tag{5.2.9}$$

$$\boldsymbol{U}^{*c} = \boldsymbol{\Delta}_{r_+^*}^{-1}(\boldsymbol{\Lambda}_{++}^{*c} + \boldsymbol{\Lambda}_{+-}^{*c}\boldsymbol{\alpha}_{-+}^{*c}), \tag{5.2.10}$$

which may be explicitly computed via the iterative algorithm of Theorem 2.35 or via any algorithm from Bean et al. (2005). Then we have the following.

Theorem 5.6 *If F corresponds to $\mathrm{PH}(\boldsymbol{\kappa}, \boldsymbol{K})$, then*

$$\bar{\boldsymbol{\Psi}}(F) = (\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}_{-+}^{*c}(\boldsymbol{I}\otimes\boldsymbol{e}). \tag{5.2.11}$$

Proof. Consider the process $\{(V_t^{*c}, J_t^{*c})\}_{t\geq 0}$ constructed above with upcrossing probabilities matrix $\boldsymbol{\alpha}_{-+}^{*c}$. Let

$$\tau^{*c} = \inf\{t > 0 : V_t^{*c} > 0, J_t^{*c} \in \mathfrak{E}_+^*\}, \tag{5.2.12}$$

and let $\{J_t^{*c}\}_{t\geq 0} =: \{(j_t^1, j_t^2)\}_{t\geq 0}$. Then, for all $i \in \mathfrak{E}_-$ and $j \in \mathfrak{E}_+$,

$$\begin{aligned}
\bar{\psi}_{ij}(F) &= \sum_{k_2=1}^{\ell}\sum_{k_1=1}^{\ell} \mathbb{P}(\tau^{*c} < \infty, j_0^2 = k_1, i_{\tau_L} = j, j_{\tau^{*c}}^2 = k_2 \mid V_0 = 0, j_0^1 = i) \\
&= \sum_{k_2=1}^{\ell}\sum_{k_1=1}^{\ell} \kappa_{k_1}\left(\boldsymbol{\alpha}_{-+}^{*c}\right)_{((i-1)\ell+k_1,(j-1)\ell+k_2)} \\
&= \left((\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}_{-+}^{*c}(\boldsymbol{I}\otimes\boldsymbol{e})\right)_{ij},
\end{aligned}$$

so the result follows. □

Corollary 5.7 *The probability of Parisian ruin with F–distributed clocks (F corresponding to $\mathrm{PH}(\boldsymbol{\kappa}, \boldsymbol{K})$) is given by*

$$\psi_i^F(u) = \boldsymbol{e}_i'\boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u}\left(\boldsymbol{I} - (\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}_{-+}^{*c}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}_{+-}^c\right)^{-1}(\boldsymbol{e} - (\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}_{-+}^{*c}\boldsymbol{e}), \tag{5.2.13}$$

with $\quad i \in \mathfrak{E}_+, u \geq 0.$

To compute the probability of cumulative Parisian ruin we follow similar steps.

Proposition 5.8 *Let $u \geq 0$, $i \in \mathfrak{E}_+$ and $n \geq 1$. Then,*

$$\mathbb{P}\left(\sum_{j=1}^{n} h_j > L > \sum_{j=1}^{n-1} h_j \mid V_0 = u, J_0 = i\right) = \left(\boldsymbol{e}_i' \boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u} \otimes \boldsymbol{\kappa}\right) \left(\boldsymbol{\alpha}_{-+}^{*c} (\boldsymbol{\beta}_{+-}^c \otimes \boldsymbol{I})\right)^{n-1} (\boldsymbol{e} - \boldsymbol{\alpha}_{-+}^{*c} \boldsymbol{e}). \tag{5.2.14}$$

Proof. For $n = 1$, the event $\{\sum_{j=1}^{n} h_j > L > \sum_{j=1}^{n-1} h_j\} = \{L < h_1\}$ implies that $\{R_t\}_{t\geq 0}$ downcrossed level 0 and it was not able to upcross it before L rang. Conditioning on the state of the first downcrossing of level 0 and on the phase–state at which L initiates, we get that

$$\begin{aligned}\mathbb{P}\left(L < h_1 \mid V_0 = u, J_0 = i\right) &= \sum_{j \in \mathfrak{E}_-, k \in \mathfrak{E}_\ell} (\boldsymbol{e}_i' \boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u})_j (\boldsymbol{\kappa}_k)(1 - (\boldsymbol{e}_j' \otimes \boldsymbol{e}_k')\boldsymbol{\alpha}_{-+}^{*c} \boldsymbol{e}) \\ &= (\boldsymbol{e}_i' \boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u} \otimes \boldsymbol{\kappa})(\boldsymbol{e} - \boldsymbol{\alpha}_{-+}^{*c} \boldsymbol{e})\end{aligned}$$

For $n \geq 2$, the event $\{\sum_{j=1}^{n} h_j > L > \sum_{j=1}^{n-1} h_j\}$ implies that $\{R_t\}_{t\geq 0}$:

1. Downcrossed 0 while in some state in $\mathfrak{E}_-$ and L initiates in some state in $\mathfrak{E}_\ell$, which happens according to the probability vector $\boldsymbol{e}_i' \boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u} \otimes \boldsymbol{\kappa}$ (ordered lexicographically);
2. For $j \in \{1, \ldots, n-1\}$,
   (a) Uprossed 0 while in some state in $\mathfrak{E}_+ \times \mathfrak{E}_\ell$ before the clock L rang, which happens according to the probability matrix $\boldsymbol{\alpha}_{-+}^{*c}$;
   (b) The clock L remained frozen in the state it was at its previous upcrossing until the next downcrossing of level 0, which happens according probability matrix $\boldsymbol{\beta}_{+-}^c \otimes \boldsymbol{I}$;
3. Did not upcross level 0 before L rang, which happens according to the probability vector $\boldsymbol{e} - \boldsymbol{\alpha}_{-+}^{*c} \boldsymbol{e}$.

Thus, (5.2.14) follows. □

Theorem 5.9 *The probability of cumulative Parisian ruin with F–distributed clock (F corresponding to $\mathrm{PH}(\boldsymbol{\kappa}, \boldsymbol{K})$) is given by*

$$\phi_i^F(u) = \left(\boldsymbol{e}_i' \boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u} \otimes \boldsymbol{\kappa}\right) \left(\boldsymbol{I} - \boldsymbol{\alpha}_{-+}^{*c} (\boldsymbol{\beta}_{+-}^c \otimes \boldsymbol{I})\right)^{-1} (\boldsymbol{e} - \boldsymbol{\alpha}_{-+}^{*c} \boldsymbol{e}), \quad i \in \mathfrak{E}_+, u \geq 0. \tag{5.2.15}$$

PROOF. Simply notice that

$$\begin{aligned}\phi_i^F(u) &= \sum_{n=1}^{\infty} \mathbb{P}\left(\sum_{j=1}^{n} h_j > L > \sum_{j=1}^{n-1} h_j \mid V_0 = u, J_0 = i\right) \\ &= \sum_{n=1}^{\infty} \left(\boldsymbol{e}_i' \boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u} \otimes \boldsymbol{\kappa}\right) \left(\boldsymbol{\alpha}_{-+}^{*c} (\boldsymbol{\beta}_{+-}^c \otimes \boldsymbol{I})\right)^{n-1} (\boldsymbol{e} - \boldsymbol{\alpha}_{-+}^{*c} \boldsymbol{e}) \\ &= \left(\boldsymbol{e}_i' \boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u} \otimes \boldsymbol{\kappa}\right) \left(\boldsymbol{I} - \boldsymbol{\alpha}_{-+}^{*c} (\boldsymbol{\beta}_{+-}^c \otimes \boldsymbol{I})\right)^{-1} (\boldsymbol{e} - \boldsymbol{\alpha}_{-+}^{*c} \boldsymbol{e}).\end{aligned}$$

The positive drift assumption of $\{V_t\}_{t\geq 0}$ implies that $\left(\boldsymbol{\beta}_{+-}^c \boldsymbol{e}\right)_j < 1$ for all $j \in \mathfrak{E}_+$. This implies that $\boldsymbol{\alpha}_{-+}^{*c}(\boldsymbol{\beta}_{+-}^c \otimes \boldsymbol{I})$ is a sub–transition matrix, $\boldsymbol{\alpha}_{-+}^{*c}(\boldsymbol{\beta}_{+-}^c \otimes \boldsymbol{I}) - \boldsymbol{I}$ is a sub–intensity matrix, and Theorem 2.8 implies that $(\boldsymbol{I} - \boldsymbol{\alpha}_{-+}^{*c}(\boldsymbol{\beta}_{+-}^c \otimes \boldsymbol{I}))^{-1}$ indeed exists. □

5.3 General fluid flow risk process

In this section we extend the results of Section 5.2 for risk processes with Brownian components. That is, we compute the probability of ruin in an infinite–horizon, Parisian and cumulative Parisian for general fluid flow risk processes.

5.3.1 Infinite–horizon probability of ruin

Let $\{R_t\}_{t\geq 0}$ be a fluid flow risk process and let $\{(V_t, J_t)\}$ be its underlying fluid flow process with $\mathfrak{E}_\sigma \neq \emptyset$. This means that the intensity matrix of $\{J_t\}_{t\geq 0}$ can be written on the form

$$\boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{\Lambda}_{\sigma\sigma} & \boldsymbol{\Lambda}_{\sigma+} & \boldsymbol{\Lambda}_{\sigma-} & \boldsymbol{\Lambda}_{\sigma 0} \\ \boldsymbol{\Lambda}_{+\sigma} & \boldsymbol{\Lambda}_{++} & \boldsymbol{\Lambda}_{+-} & \boldsymbol{\Lambda}_{+0} \\ \boldsymbol{\Lambda}_{-\sigma} & \boldsymbol{\Lambda}_{-+} & \boldsymbol{\Lambda}_{--} & \boldsymbol{\Lambda}_{-0} \\ \boldsymbol{\Lambda}_{0\sigma} & \boldsymbol{\Lambda}_{0+} & \boldsymbol{\Lambda}_{0-} & \boldsymbol{\Lambda}_{00} \end{pmatrix},$$

and the reward rates $\{r_i\}_{i\in\mathfrak{E}_\sigma\cup\mathfrak{E}_+\cup\mathfrak{E}_-\cup\mathfrak{E}_0}$ and Brownian noise rates $\{\sigma_i\}_{i\in\mathfrak{E}_\sigma\cup\mathfrak{E}_+\cup\mathfrak{E}_-\cup\mathfrak{E}_0}$ are such that

$$\begin{aligned} r_i \in \mathbb{R} \quad &\text{and} \quad \sigma_i > 0 \quad \text{for all} \quad i \in \mathfrak{E}_\sigma, \\ r_i > 0 \quad &\text{and} \quad \sigma_i = 0 \quad \text{for all} \quad i \in \mathfrak{E}_+, \\ r_i < 0 \quad &\text{and} \quad \sigma_i = 0 \quad \text{for all} \quad i \in \mathfrak{E}_-, \\ r_i = 0 \quad &\text{and} \quad \sigma_i = 0 \quad \text{for all} \quad i \in \mathfrak{E}_0. \end{aligned}$$

Let $\boldsymbol{\Lambda}^c$ correspond to the generator of the $\mathfrak{E}_0$–censored process $\{J_t\}_{t\geq 0}$, so that

$$\boldsymbol{\Lambda}^c = \begin{pmatrix} \boldsymbol{\Lambda}^c_{\sigma\sigma} & \boldsymbol{\Lambda}^c_{\sigma+} & \boldsymbol{\Lambda}^c_{\sigma-} \\ \boldsymbol{\Lambda}^c_{+\sigma} & \boldsymbol{\Lambda}^c_{++} & \boldsymbol{\Lambda}^c_{+-} \\ \boldsymbol{\Lambda}^c_{-\sigma} & \boldsymbol{\Lambda}_{-+}c & \boldsymbol{\Lambda}^c_{--} \end{pmatrix}$$

as in (2.4.17). By Theorem 2.31, Theorem 2.33 and Subsection 2.4.2, we know that if $\tau_0 = \inf\{t > 0 : V_t < 0\}$ and $m_0 = J_{\gamma_0}$,

$$\mathbb{P}(m_0 = j, \tau_0 < \infty \mid J_0 = i) = \left(\boldsymbol{e}_i' \boldsymbol{\beta}^c e^{\boldsymbol{D}^c u}\right)_j, \quad i \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+, j \in \mathfrak{E}_\sigma \cup \mathfrak{E}_-,$$

where

$$\boldsymbol{\beta}^c = \begin{pmatrix} \boldsymbol{\beta}^c_{\sigma d} \\ \boldsymbol{\beta}^c_{+d} \end{pmatrix} := \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{\beta}^c_{+\sigma} & \boldsymbol{\beta}^c_{+-} \end{pmatrix} =: \begin{pmatrix} \boldsymbol{\beta}^c_{u\sigma} & \boldsymbol{\beta}^c_{u-} \end{pmatrix}, \quad \text{and}$$
$$\boldsymbol{D}^c = \begin{pmatrix} \boldsymbol{D}^c_{\sigma\sigma} & \boldsymbol{D}^c_{\sigma-} \\ \boldsymbol{D}^c_{-\sigma} & \boldsymbol{D}^c_{--} \end{pmatrix} =: \begin{pmatrix} \boldsymbol{D}^c_{\sigma d} \\ \boldsymbol{D}^c_{-d} \end{pmatrix}$$

satisfy the equations

$$\begin{aligned} &\boldsymbol{\Delta}^{-1}_{r_+} \boldsymbol{\Lambda}^c_{+d} + \boldsymbol{\Delta}^{-1}_{r_+} \boldsymbol{\Lambda}^c_{++} \boldsymbol{\beta}^c_{+d} + \boldsymbol{\beta}^c_{+d} \boldsymbol{D}^c = \boldsymbol{0}, \\ &\boldsymbol{D}^c_{-d} = \boldsymbol{\Lambda}^c_{-d} + \boldsymbol{\Lambda}^c_{-+} \boldsymbol{\beta}^c_{+d}, \\ &\boldsymbol{D}^c_{\sigma d} \boldsymbol{D}^c + 2\boldsymbol{\Delta}^{-2}_{\sigma} (\boldsymbol{\Delta}_{r_\sigma} \boldsymbol{D}^c_{\sigma d} + \boldsymbol{\Lambda}^c_{\sigma d} + \boldsymbol{\Lambda}^c_{\sigma+} \boldsymbol{\beta}^c_{+d}) = \boldsymbol{0}, \end{aligned}$$

and may be computed via the iterative algorithm of Theorem 2.32. Since $\{R_t\}_{t\geq 0}$ and $\{V_t\}_{t\geq 0}$ share the same first passage probabilities, we have the following.

Theorem 5.10 *The infinite–horizon probability of ruin* $\psi_i(u) := \mathbb{P}(\inf_t R_t < 0 \mid R_0 = 0, J_0 = i)$ *is given by*

$$\psi_i(u) = \boldsymbol{e}_i' \boldsymbol{\beta}^c e^{\boldsymbol{D}^c u} \boldsymbol{e}, \quad i \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+, u \geq 0.$$

5.3.2 Parisian and cumulative Parisian ruin

The case of Parisian ruin for a fluid flow process with a genuine Brownian component is slightly more complicated than it non–Brownian counterpart. Since $\mathfrak{E}_\sigma \neq \emptyset$, if the fluid flow process is zero at time t while $J_t = i \in \mathfrak{E}_\sigma$, then any compact interval $[t, t+\delta]$, $\delta > 0$, contains an infinite number of points where $\{R_t\}_{t\geq 0}$ is zero again (see Section 10 in Rogers and Williams (1993)). The precise definition of Parisian ruin for general fluid flow risk processes is the following.

Definition 5.11 (Parisian ruin) For a fluid flow risk process $\{R_t\}_{t\geq 0}$, let G denote the set of left–end points of the subexcursions below 0 of $\{R_t\}_{t\geq 0}$, and let $f : G \to \mathbb{R}_+$ be the function that maps each left–end point of a subexcursion below 0 with its right–end point. Let $\{L_g\}_{g\in G}$ be a collection of i.d.d.r.v. with common distribution F, independent from $\{R_t\}_{t\geq 0}$. Then we say that $\{R_t\}_{t\geq 0}$ gets ruined in a **Parisian** way with F–distributed clocks if there exists $g_0 \in G$ such that $L_{g_0} < f(g_0) - g_0$. For any $i \in \mathfrak{E}_+ \cup \mathfrak{E}_\sigma$ and $u \geq 0$, we define the probability of Parisian ruin $\psi_i^F(u)$ by

$$\psi_i^F(u) = \mathbb{P}\left(\cup_{g\in G}\{L_g < f(g) - g\} \mid J_0 = i, R_0 = u\right). \tag{5.3.1}$$

The definition of cumulative Parisian ruin for general fluid flow risk processes is stated next.

Definition 5.12 (Cumulative Parisian ruin) Let $\{R_t\}_{t\geq 0}$ be fluid flow risk process and let L be an independent random variable with distribution F. Then we say that $\{R_t\}_{t\geq 0}$ gets ruined in a **cumulative Parisian** way with F–distributed clock if $\int_0^\infty \mathbb{1}\{R_s < 0\}\mathrm{d}s > L$. For any $i \in \mathfrak{E}_+ \cup \mathfrak{E}_\sigma$ and $u \geq 0$, we define the probability of cumulative Parisian ruin $\phi_i^F(u)$ by

$$\phi_i^F(u) = \mathbb{P}\left(\int_0^\infty \mathbb{1}\{R_s < 0\}\mathrm{d}s > L \mid J_0 = i, R_0 = u\right). \tag{5.3.2}$$

With the framework we have presented in this thesis, it is not possible to attack the problem of computing (5.3.1) or (5.3.2) directly. To avoid this complication, we follow a standard trick based on approximations (see e.g. Loeffen et al. (2013) or Baurdoux et al. (2016)), which we describe next.

Fix $\epsilon > 0$, let $u_0^\epsilon = 0$, and for $n \geq 1$ define

$$\begin{aligned} d_n^\epsilon &= \inf\{t \geq u_{n-1}^\epsilon : V_t < -\epsilon\} \\ u_n^\epsilon &= \inf\{t > d_n^\epsilon : V_t \geq 0\}, \end{aligned}$$

so that $\{d_n^\epsilon\}_{n\geq 1}$ and $\{u_n^\epsilon\}_{n\geq 1}$ correspond to the alternating points in time in which $\{V_t\}_{t\geq 0}$ downcrosses level $-\epsilon$, later to upcross level 0, respectively. For all $n \geq 1$, let

$$h_n^\epsilon = \int_{d_n^\epsilon}^{u_n^\epsilon} \mathbb{1}\{J_s \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+ \cup \mathfrak{E}_0\}\mathrm{d}s, \tag{5.3.3}$$

with the convention that $h_n^\epsilon = 0$ for $d_n^\epsilon = \infty$. Then h_n^ϵ corresponds to total amount of actual time, in the sense of the equivalent process $\{R_t\}_{t\geq 0}$, accumulated in the interval $[d_n^\epsilon, u_n^\epsilon)$. See Figure 5.3 for a visual description of $\{h_n^\epsilon\}_{n\geq 1}$.

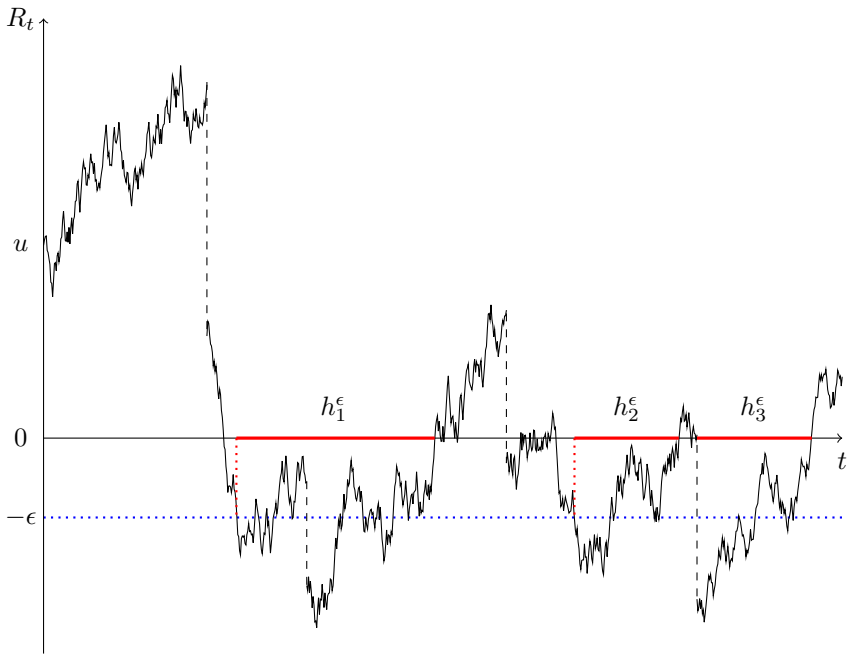


Figure 5.3: A fluid flow risk processes with $\mathfrak{E}_\sigma \neq \emptyset$. For given $\epsilon > 0$, the intervals in red correspond to the succesive times at which $\{R_t\}_{t\geq 0}$ downcrosses $-\epsilon$ and upcrosses 0 afterwards. The lengths of these intervals correspond to h_i^ϵ, $i \geq 1$.

Definition 5.13 (ϵ–Parisian ruin) Let $\{R_t\}_{t\geq 0}$ be fluid flow risk process, let $\{h_n^\epsilon\}_{n\geq 1}$ be defined as in (5.3.3), and let $\{L_n\}_{n\geq 1}$ be an independent i.i.d. sequence of random variables with common distribution F. We say that $\{R_t\}_{t\geq 0}$ gets ruined in a ϵ–**Parisian** way with F–distributed clocks if there exists some $n \geq 1$ such that $L_n < h_n^\epsilon$. For any $i \in \mathfrak{E}_+ \cup \mathfrak{E}_\sigma$ and $u \geq 0$, we define the probability of ϵ–Parisian ruin $\psi_i^F(u, \epsilon)$ by

$$\psi_i^F(u, \epsilon) = \mathbb{P}\left(\cup_{n\geq 1}\{L_n < h_n^\epsilon\} \mid J_0 = i, R_0 = u\right). \tag{5.3.4}$$

Definition 5.14 (ϵ–Cumulative Parisian ruin) Let $\{R_t\}_{t\geq 0}$ be fluid flow risk process, let $\{h_n^\epsilon\}_{n\geq 1}$ be defined as in (5.3.3) and let L be an independent random variable with common distribution F. We say that $\{R_t\}_{t\geq 0}$ gets ruined in a ϵ–**cumulative Parisian** way with F–distributed clock if such that $L < \sum_{n\geq 1} h_n^\epsilon$. For any $i \in \mathfrak{E}_+ \cup \mathfrak{E}_\sigma$ and $u \geq 0$, we define the probability of ϵ–

cumulative Parisian ruin $\phi_i^F(u,\epsilon)$ by

$$\phi_i^F(u,\epsilon) = \mathbb{P}\left(L < \sum_{n=1}^{\infty} h_n^\epsilon \mid J_0 = i, R_0 = u\right). \tag{5.3.5}$$

The limits

$$\lim_{\epsilon\downarrow 0} \psi_i^F(u,\epsilon), \quad \text{and} \quad \lim_{\epsilon\downarrow 0} \phi_i^F(u,\epsilon) \tag{5.3.6}$$

clearly exist. Indeed, a pathwise inspection reveals that $\psi_i^F(u,\epsilon_1) \leq \psi_i^F(u,\epsilon_2)$ and $\phi_i^F(u,\epsilon_1) \leq \phi_i^F(u,\epsilon_2)$ if $\epsilon_1 > \epsilon_2 > 0$, so that $\psi_i^F(u,\cdot)$ and $\phi_i^F(u,\cdot)$ are monotonous. In fact, following the steps of Section 1.2 in Baurdoux et al. (2016), it is straightforward to check that

$$\psi_i^F(u) = \lim_{\epsilon\downarrow 0} \psi_i^F(u,\epsilon), \quad \text{and} \quad \phi_i^F(u) = \lim_{\epsilon\downarrow 0} \phi_i^F(u,\epsilon).$$

Thus, the study of the both Parisian and cumulative Parisian ruin for the general fluid flow model requires two steps: computing the probabilities of ϵ–Parisian and ϵ–cumulative Parisian ruin, and then taking $\epsilon \to 0$. The first step is very similar to the analysis made in Subsection 5.2.2, which we carry on next.

Conditional on $V_0 = -\epsilon$ and $J_0 = i \in \mathfrak{E}_\sigma \cup \mathfrak{E}_-$, define

$$\tau = \inf\{t > 0 : V_t = 0, J_t \notin \mathfrak{E}_-\} \text{ and } h = \int_0^\tau \mathbb{1}\{V_s < 0, J_s \notin \mathfrak{E}_-\}\mathrm{d}s.$$

Let $L \sim F$ be independent of $\{(V_t, J_t)\}_{t\geq 0}$, and for $i \in \mathfrak{E}_\sigma \cup \mathfrak{E}_-$, $j \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+$ let

$$\bar{\psi}_{ij}(F,\epsilon) = \mathbb{P}(h < L, J_{\tau_L} = j \mid V_0 = -\epsilon, J_0 = i).$$

Finally, let $\bar{\boldsymbol{\Psi}}(F,\epsilon) = \{\bar{\psi}_{ij}(F,\epsilon)\}_{i\in\mathfrak{E}_\sigma\cup\mathfrak{E}_-, j\in\mathfrak{E}_\sigma\cup\mathfrak{E}_+}$ denote the corresponding matrix. Then we have the following result.

Proposition 5.15 *Let $u \geq 0$, $i \in \mathfrak{E}_+$ and $n \geq 1$. Then,*

$$\begin{aligned}&\mathbb{P}\left(\bigcap_{j<n}\{L_j > h_j^\epsilon\}, L_n < h_n^\epsilon \mid V_0 = u, J_0 = i\right)\\ &= \boldsymbol{e}_i'\boldsymbol{\beta}^c e^{\boldsymbol{D}^c(u+\epsilon)}\left(\bar{\boldsymbol{\Psi}}(F,\epsilon)\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon}\right)^{n-1}(\boldsymbol{e} - \bar{\boldsymbol{\Psi}}(F,\epsilon)\boldsymbol{e}).\end{aligned} \tag{5.3.7}$$

Proof. Follows along the same lines as the proof of Proposition 5.4. □

Theorem 5.16 *For any $i \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+$ and $u \geq 0$, the probability of ϵ–Parisian ruin $\psi_i^F(u)$ is given by*

$$\psi_i^F(u,\epsilon) = \boldsymbol{e}_i'\boldsymbol{\beta}^c e^{\boldsymbol{D}^c(u+\epsilon)}\left(\boldsymbol{I} - \bar{\boldsymbol{\Psi}}(F,\epsilon)\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon}\right)^{-1}(\boldsymbol{e} - \bar{\boldsymbol{\Psi}}(L,\epsilon)\boldsymbol{e}).$$

Proof. We have that

$$\begin{aligned}\psi_i^F(u,\epsilon) &= \sum_{n=1}^{\infty}\mathbb{P}\left(\cap_{j<n}\{L_j > h_j^\epsilon\}, L_n < h_n^\epsilon \mid V_0 = u, J_0 = i\right)\\ &= \sum_{n=1}^{\infty}\boldsymbol{e}_i'\boldsymbol{\beta}^c e^{\boldsymbol{D}^c(u+\epsilon)}\left(\bar{\boldsymbol{\Psi}}(F,\epsilon)\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon}\right)^{n-1}(\boldsymbol{e} - \bar{\boldsymbol{\Psi}}(F,\epsilon)\boldsymbol{e})\\ &= \boldsymbol{e}_i'\boldsymbol{\beta}^c e^{\boldsymbol{D}^c(u+\epsilon)}\left(\boldsymbol{I} - \bar{\boldsymbol{\Psi}}(F,\epsilon)\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon}\right)^{-1}(\boldsymbol{e} - \bar{\boldsymbol{\Psi}}(F,\epsilon)\boldsymbol{e}). \qquad (5.3.8)\end{aligned}$$

As in the proof of Theorem 5.5, one can argue that $V_t \to \infty$ as $t \to \infty$ implies that

$$\left(\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon}\boldsymbol{e}\right)_j < 1 \quad \text{for all } j \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+ \text{ and } \epsilon > 0, \qquad (5.3.9)$$

so that $\boldsymbol{I} - \bar{\boldsymbol{\Psi}}(F,\epsilon)\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon}$ in (5.3.8) is indeed invertible. □ Just as

in Subsection 5.2.2, $\bar{\boldsymbol{\Psi}}(F,\epsilon)$ will be explicitly computable given the assumption that F corresponds to a phase–type distribution.

Theorem 5.17 *If F corresponds to* $\mathrm{PH}(\boldsymbol{\kappa},\boldsymbol{K})$*, then*

$$\bar{\boldsymbol{\Psi}}(F,\epsilon) = (\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}e^{\boldsymbol{U}_{*c}\epsilon}(\boldsymbol{I}\otimes\boldsymbol{e}) \qquad (5.3.10)$$

where

$$\boldsymbol{\alpha}^{*c} = \begin{pmatrix}\boldsymbol{\alpha}^{*c}_{\sigma u}\\ \boldsymbol{\alpha}^{*c}_{-u}\end{pmatrix} = \begin{pmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{\alpha}^{*c}_{-\sigma} & \boldsymbol{\alpha}^{*c}_{-+}\end{pmatrix}, \quad \textit{and}$$
$$\boldsymbol{U}^{*c} = \begin{pmatrix}\boldsymbol{U}^{*c}_{\sigma u}\\ \boldsymbol{U}^{*c}_{+u}\end{pmatrix} = \begin{pmatrix}\boldsymbol{U}^{*c}_{\sigma\sigma} & \boldsymbol{U}^{*c}_{\sigma+}\\ \boldsymbol{U}^{*c}_{+\sigma} & \boldsymbol{U}^{*c}_{++}\end{pmatrix}$$

satisfy the equations

$$\begin{aligned}&\boldsymbol{\Lambda}^{*c}_{-u} + \boldsymbol{\Lambda}^{*c}_{--}\boldsymbol{\alpha}^{*c}_{-u} + \boldsymbol{\alpha}^{*c}_{-u}\boldsymbol{U}^{*c} = \boldsymbol{0},\\ &\boldsymbol{U}^{*c}_{+u} = \boldsymbol{\Delta}^{-1}_{r^*_+}(\boldsymbol{\Lambda}^{*c}_{+u} + \boldsymbol{\Lambda}^{*c}_{+-}\boldsymbol{\alpha}^{*c}_{-u}),\\ &\boldsymbol{U}^{*c}_{\sigma u}\boldsymbol{U}^{*c} + 2\boldsymbol{\Delta}^{-2}_{\sigma^*}(-\boldsymbol{\Delta}_{r^*_\sigma}\boldsymbol{U}^{*c}_{\sigma u} + \boldsymbol{\Lambda}^{*c}_{\sigma u} + \boldsymbol{\Lambda}^{*c}_{\sigma-}\boldsymbol{\alpha}^{*c}_{-u}) = \boldsymbol{0},\end{aligned}$$

and whose exact value may be computed via the iterative algorithm of Theorem 2.35.

PROOF. Follows along the same lines as the proof of Theorem 5.6 with the extra set of Brownian states. □

Corollary 5.18 *The probability of ϵ–Parisian ruin with F–distributed clocks (F corresponding to $\mathrm{PH}(\boldsymbol{\kappa}, \boldsymbol{K})$) is given by*

$$\psi_i^F(u,\epsilon) = \boldsymbol{e}_i'\boldsymbol{\beta}^c e^{\boldsymbol{D}^c(u+\epsilon)}\left(\boldsymbol{I} - (\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}e^{\boldsymbol{U}^{*c}\epsilon}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon}\right)^{-1} \times (\boldsymbol{e} - (\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e})$$

for $i \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+, u \geq 0$.

Concerning ϵ–cumulative Parisian ruin, we have the following.

Proposition 5.19 *Let $u \geq 0$, $i \in \mathfrak{E}_+$ and $n \geq 1$. Then,*

$$\mathbb{P}\left(\sum_{j=1}^{n} h_j^\epsilon > L > \sum_{j=1}^{n-1} h_j^\epsilon \mid V_0 = u, J_0 = i\right) = \left(\boldsymbol{e}_i'\boldsymbol{\beta}^c e^{\boldsymbol{D}^c(u+\epsilon)}\otimes\boldsymbol{\kappa}\right)\left(\boldsymbol{\alpha}^{*c}e^{\boldsymbol{U}^{*c}\epsilon}[(\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon})\otimes\boldsymbol{I}]\right)^{n-1}(\boldsymbol{e} - \boldsymbol{\alpha}_{-+}^{*c}e^{\boldsymbol{U}^{*c}}\boldsymbol{e}). \tag{5.3.11}$$

PROOF. Follows along the same lines of the proof of Proposition 5.8.

Theorem 5.20 *The probability of ϵ–cumulative Parisian ruin with F–distributed clock (F corresponding to $\mathrm{PH}(\boldsymbol{\kappa}, \boldsymbol{K})$) is given by*

$$\phi_i^F(u,\epsilon) = \left(\boldsymbol{e}_i'\boldsymbol{\beta}^c e^{\boldsymbol{D}^c(u+\epsilon)}\otimes\boldsymbol{\kappa}\right)\left(\boldsymbol{I} - \boldsymbol{\alpha}^{*c}e^{\boldsymbol{U}^{*c}\epsilon}[(\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon})\otimes\boldsymbol{I}]\right)^{-1}(\boldsymbol{e} - \boldsymbol{\alpha}^{*c}e^{\boldsymbol{U}^{*c}}\boldsymbol{e})$$

for $i \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+$ and $u \geq 0$.

PROOF. Notice that

$$\begin{aligned}\phi_i^F(u,\epsilon) &= \sum_{n=1}^{\infty}\mathbb{P}\left(\sum_{j=1}^{n} h_j^\epsilon > L > \sum_{j=1}^{n-1} h_j^\epsilon \mid V_0 = u, J_0 = i\right)\\ &= \sum_{n=1}^{\infty}\left(\boldsymbol{e}_i'\boldsymbol{\beta}^c e^{\boldsymbol{D}^c(u+\epsilon)}\otimes\boldsymbol{\kappa}\right)\left(\boldsymbol{\alpha}^{*c}e^{\boldsymbol{U}^{*c}\epsilon}[(\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon})\otimes\boldsymbol{I}]\right)^{n-1}(\boldsymbol{e} - \boldsymbol{\alpha}_{-+}^{*c}e^{\boldsymbol{U}^{*c}}\boldsymbol{e})\\ &= \left(\boldsymbol{e}_i'\boldsymbol{\beta}^c e^{\boldsymbol{D}^c(u+\epsilon)}\otimes\boldsymbol{\kappa}\right)\left(\boldsymbol{I} - \boldsymbol{\alpha}^{*c}e^{\boldsymbol{U}^{*c}\epsilon}[(\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon})\otimes\boldsymbol{I}]\right)^{-1}(\boldsymbol{e} - \boldsymbol{\alpha}^{*c}e^{\boldsymbol{U}^{*c}}\boldsymbol{e}).\end{aligned}$$

The positive drift assumption of $\{V_t\}_{t\geq 0}$ implies that $\left(\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon})\boldsymbol{e}\right)_j < 1$ for all $j \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+$. This implies that $\boldsymbol{\alpha}^{*c} e^{\boldsymbol{U}^{*c}\epsilon}[(\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon}) \otimes \boldsymbol{I}]$ is a sub–transition matrix, $\boldsymbol{\alpha}^{*c} e^{\boldsymbol{U}^{*c}\epsilon}[(\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon}) \otimes \boldsymbol{I}] - \boldsymbol{I}$ is a sub–intensity matrix, and Theorem 2.8 implies that $(\boldsymbol{I} - \boldsymbol{\alpha}^{*c} e^{\boldsymbol{U}^{*c}\epsilon}[(\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon}) \otimes \boldsymbol{I}])^{-1}$ indeed exists. □

As $\epsilon \downarrow 0$ we get the following results.

Theorem 5.21 *In the case $\mathfrak{E}_\sigma \neq \emptyset$ and $F \sim \mathrm{PH}(\boldsymbol{\kappa}, \boldsymbol{K})$, the probability of Parisian ruin is given by*

$$\begin{aligned}\psi_i^F(u) &= \boldsymbol{e}_i' \boldsymbol{\beta}^c e^{\boldsymbol{D}^c u} \\ &\quad \times \begin{pmatrix} -(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{U}_{\sigma u}^{*c}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}_{u\sigma}^c - \boldsymbol{D}_{\sigma\sigma}^c & -(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{U}_{\sigma+}^{*c}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}_{+-}^c - \boldsymbol{D}_{\sigma-}^c \\ -(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}_{-u}^{*c}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}_{u\sigma}^c & -(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}_{-+}^{*c}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}_{+-}^c + \boldsymbol{I} \end{pmatrix}^{-1} \\ &\quad \times \begin{pmatrix} -(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{U}_{\sigma u}^{*c}\boldsymbol{e} \\ \boldsymbol{e} - (\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}_{-u}^{*c}\boldsymbol{e} \end{pmatrix}, \quad i \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+, u \geq 0, \end{aligned} \tag{5.3.12}$$

and the probability of cumulative Parisian ruin is given by

$$\phi_i^F(u) = \left(\boldsymbol{e}_i' \boldsymbol{\beta}^c e^{\boldsymbol{D}^c u} \otimes \boldsymbol{\kappa}\right) \tag{5.3.13}$$

$$\begin{aligned} &\times \begin{pmatrix} -\boldsymbol{U}_{\sigma u}^{*c}(\boldsymbol{\beta}_{u\sigma}^c \otimes I) - \boldsymbol{D}_{\sigma\sigma}^c \otimes \boldsymbol{I} & -\boldsymbol{U}_{\sigma+}^{*c}(\boldsymbol{\beta}_{+-}^c \otimes I) - \boldsymbol{D}_{\sigma-}^c \otimes \boldsymbol{I} \\ -\boldsymbol{\alpha}_{-u}^{*c}(\boldsymbol{\beta}_{u\sigma}^c \otimes \boldsymbol{I}) & -\boldsymbol{\alpha}_{-u}^{*c}(\boldsymbol{\beta}_{u-}^c \otimes \boldsymbol{I}) + \boldsymbol{I} \end{pmatrix}^{-1} \\ &\times \begin{pmatrix} -\boldsymbol{U}_{\sigma u}^{*c}\boldsymbol{e} \\ \boldsymbol{e} - \boldsymbol{\alpha}_{-u}^{*c}\boldsymbol{e} \end{pmatrix}, \quad i \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+, u \geq 0. \end{aligned} \tag{5.3.14}$$

Proof. We first prove (5.3.14). Notice that

$$\begin{aligned}\phi_i^F(u) &= \lim_{\epsilon\to 0} \left(\boldsymbol{e}_i' \boldsymbol{\beta}^c e^{\boldsymbol{D}^c(u+\epsilon)} \otimes \boldsymbol{\kappa}\right) \left(\boldsymbol{I} - \boldsymbol{\alpha}^{*c} e^{\boldsymbol{U}^{*c}\epsilon} \left[\left(\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon}\right) \otimes \boldsymbol{I}\right]\right)^{-1} \\ &\quad \times \left(\boldsymbol{e} - \begin{pmatrix} \boldsymbol{\alpha}_{\sigma u}^{*c} \\ \boldsymbol{\alpha}_{-u}^{*c} \end{pmatrix} e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e}\right) \\ &= \lim_{\epsilon\to 0} \left(\boldsymbol{e}_i' \boldsymbol{\beta}^c e^{\boldsymbol{D}^c(u+\epsilon)} \otimes \boldsymbol{\kappa}\right) \left(\begin{pmatrix} \epsilon^{-1}\boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \left[\boldsymbol{I} - \boldsymbol{\alpha}_{du}^{*c} e^{\boldsymbol{U}^{*c}\epsilon} \left[\left(\boldsymbol{\beta}^c e^{\boldsymbol{D}^c\epsilon}\right) \otimes \boldsymbol{I}\right]\right]\right)^{-1} \\ &\quad \times \begin{pmatrix} \epsilon^{-1}\boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \left(\boldsymbol{e} - \begin{pmatrix} \boldsymbol{\alpha}_{\sigma u}^{*c} \\ \boldsymbol{\alpha}_{-u}^{*c} \end{pmatrix} e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e}\right). \end{aligned}$$

Since

$$\begin{aligned}
&\lim_{\epsilon\to 0}\begin{pmatrix}\epsilon^{-1}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{pmatrix}\left(\boldsymbol{e}-\begin{pmatrix}\boldsymbol{\alpha}^{*c}_{\sigma u}\\ \boldsymbol{\alpha}^{*c}_{-u}\end{pmatrix}e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e}\right)\\
&\quad=\lim_{\epsilon\to 0}\begin{pmatrix}\epsilon^{-1}(\boldsymbol{e}-\boldsymbol{\alpha}^{*c}_{\sigma u}e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e})\\ \boldsymbol{e}-\boldsymbol{\alpha}^{*c}_{-u}e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e}\end{pmatrix}\\
&\quad=\lim_{\epsilon\to 0}\begin{pmatrix}\epsilon^{-1}(\boldsymbol{e}-\boldsymbol{\alpha}^{*c}_{\sigma u}(\boldsymbol{I}+\boldsymbol{U}^{*c}\epsilon+o(\epsilon))\boldsymbol{e})\\ \boldsymbol{e}-\boldsymbol{\alpha}^{*c}_{-u}e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e}\end{pmatrix}\\
&\quad=\lim_{\epsilon\to 0}\begin{pmatrix}\epsilon^{-1}(\boldsymbol{e}-\begin{pmatrix}\boldsymbol{I} & \boldsymbol{0}\end{pmatrix}(\boldsymbol{I}+\boldsymbol{U}^{*c}\epsilon+o(\epsilon))\boldsymbol{e})\\ \boldsymbol{e}-\boldsymbol{\alpha}^{*c}_{-u}e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e}\end{pmatrix}=\begin{pmatrix}-\boldsymbol{U}^{*c}_{\sigma u}\boldsymbol{e}\\ \boldsymbol{e}-\boldsymbol{\alpha}^{*c}_{-u}\boldsymbol{e}\end{pmatrix}
\end{aligned}$$

and

$$\begin{aligned}
&\lim_{\epsilon\to 0}\begin{pmatrix}\epsilon^{-1}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{pmatrix}\left[\boldsymbol{I}-\boldsymbol{\alpha}^{*c}e^{\boldsymbol{U}^{*c}\epsilon}\left[\left(\boldsymbol{\beta}^{c}e^{\boldsymbol{D}^{c}\epsilon}\right)\otimes\boldsymbol{I}\right]\right]\\
&\quad=\lim_{\epsilon\to 0}\begin{pmatrix}\epsilon^{-1}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{pmatrix}\left[\boldsymbol{I}-\boldsymbol{\alpha}^{*c}\left(\boldsymbol{I}+\boldsymbol{U}^{*c}\epsilon+o(\epsilon)\right)\left[\left(\boldsymbol{\beta}^{c}\left(\boldsymbol{I}+\boldsymbol{D}^{c}\epsilon+o(\epsilon)\right)\right)\otimes\boldsymbol{I}\right]\right]\\
&\quad=\lim_{\epsilon\to 0}\begin{pmatrix}\epsilon^{-1}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{pmatrix}\\
&\qquad\times\left[\boldsymbol{I}-\left\{\boldsymbol{\alpha}^{*c}(\boldsymbol{\beta}^{c}\otimes\boldsymbol{I})+\boldsymbol{\alpha}^{*c}\boldsymbol{U}^{*c}(\boldsymbol{\beta}^{c}\otimes\boldsymbol{I})\epsilon+\boldsymbol{\alpha}^{*c}\left((\boldsymbol{\beta}^{c}\boldsymbol{D}^{c})\otimes\boldsymbol{I}\right)\epsilon+o(\epsilon)\right\}\right]\\
&\quad=-\begin{pmatrix}\boldsymbol{U}^{*c}_{\sigma u}(\boldsymbol{\beta}^{c}_{u\sigma}\otimes I)+\boldsymbol{D}^{c}_{\sigma\sigma}\otimes\boldsymbol{I} & \boldsymbol{U}^{*c}_{\sigma +}(\boldsymbol{\beta}^{c}_{+-}\otimes I)+\boldsymbol{D}^{c}_{\sigma -}\otimes\boldsymbol{I}\\ \boldsymbol{\alpha}^{*c}_{-u}(\boldsymbol{\beta}^{c}_{u\sigma}\otimes\boldsymbol{I}) & \boldsymbol{\alpha}^{*c}_{-u}(\boldsymbol{\beta}^{c}_{u-}\otimes\boldsymbol{I})-\boldsymbol{I}\end{pmatrix},
\end{aligned}\tag{5.3.15}$$

so that (5.3.14) follows. The nonsingularity of (5.3.15) follows by noticing that the matrix is the negative of a sub–intensity matrix.

To prove (5.3.12),

$$\begin{aligned}
\psi_i^F(u)&=\lim_{\epsilon\to 0}\boldsymbol{e}_i'\boldsymbol{\beta}^{c}e^{\boldsymbol{D}^{c}(u+\epsilon)}\left(\boldsymbol{I}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}e^{\boldsymbol{U}^{*c}\epsilon}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}^{c}e^{D^{c}\epsilon}\right)^{-1}\\
&\qquad\times\left(\boldsymbol{e}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\begin{pmatrix}\boldsymbol{\alpha}^{*c}_{\sigma u}\\ \boldsymbol{\alpha}^{*c}_{-u}\end{pmatrix}e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e}\right)\\
&=\lim_{\epsilon\to 0}\boldsymbol{e}_i'\boldsymbol{\beta}^{c}_{+-}e^{\boldsymbol{D}(u+\epsilon)}\\
&\qquad\times\left(\begin{pmatrix}\epsilon^{-1}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{pmatrix}\left[\boldsymbol{I}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}e^{\boldsymbol{U}^{*c}\epsilon}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}^{c}e^{D^{c}\epsilon}\right]\right)^{-1}\\
&\qquad\times\begin{pmatrix}\epsilon^{-1}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{pmatrix}\left(\boldsymbol{e}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\begin{pmatrix}\boldsymbol{\alpha}^{*c}_{\sigma u}\\ \boldsymbol{\alpha}^{*c}_{-u}\end{pmatrix}e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e}\right).
\end{aligned}$$

Now,

$$
\begin{aligned}
&\lim_{\epsilon\to 0}\begin{pmatrix}\epsilon^{-1}\boldsymbol{I} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{I}\end{pmatrix}\left(\boldsymbol{e}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\begin{pmatrix}\boldsymbol{\alpha}^{*c}_{\sigma u}\\ \boldsymbol{\alpha}^{*c}_{-u}\end{pmatrix}e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e}\right)\\
&\quad=\lim_{\epsilon\to 0}\left(\begin{pmatrix}\epsilon^{-1}\left(\boldsymbol{e}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}_{\sigma u}e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e}\right)\\ \boldsymbol{e}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}_{-u}e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e}\end{pmatrix}\right)\\
&\quad=\lim_{\epsilon\to 0}\left(\begin{pmatrix}\epsilon^{-1}\left(\boldsymbol{e}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}_{\sigma u}(\boldsymbol{I}+\boldsymbol{U}^{*c}\epsilon+o(\epsilon))\boldsymbol{e}\right)\\ \boldsymbol{e}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}_{-u}e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e}\end{pmatrix}\right)\\
&\quad=\lim_{\epsilon\to 0}\left(\begin{pmatrix}\epsilon^{-1}\left(\boldsymbol{e}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\begin{pmatrix}\boldsymbol{I} & \mathbf{0}\end{pmatrix}(\boldsymbol{I}+\boldsymbol{U}^{*c}\epsilon+o(\epsilon))\boldsymbol{e}\right)\\ \boldsymbol{e}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}_{-u}e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e}\end{pmatrix}\right)\\
&\quad=\lim_{\epsilon\to 0}\left(\begin{pmatrix}\epsilon^{-1}\left(\boldsymbol{e}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{e}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{U}^{*c}_{\sigma u}\boldsymbol{e}\epsilon+o(\epsilon)\right)\\ \boldsymbol{e}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}_{-u}e^{\boldsymbol{U}^{*c}\epsilon}\boldsymbol{e}\end{pmatrix}\right)\\
&\quad=\begin{pmatrix}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{U}^{*c}_{\sigma u}\boldsymbol{e}\\ \boldsymbol{e}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}_{-u}\boldsymbol{e}\end{pmatrix}.
\end{aligned}
$$

Additionally, we have that

$$
\begin{aligned}
&\lim_{\epsilon\to 0}\begin{pmatrix}\epsilon^{-1}\boldsymbol{I} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{I}\end{pmatrix}\left[\boldsymbol{I}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}e^{\boldsymbol{U}^{*c}\epsilon}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}^{c}e^{\boldsymbol{D}^{c}\epsilon}\right]\\
&\quad=\lim_{\epsilon\to 0}\begin{pmatrix}\epsilon^{-1}\boldsymbol{I} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{I}\end{pmatrix}[\boldsymbol{I}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}(\boldsymbol{I}+\boldsymbol{U}^{*c}\epsilon+o(\epsilon))(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}^{c}(\boldsymbol{I}+\boldsymbol{D}^{c}\epsilon+o(\epsilon))]\\
&\quad=\lim_{\epsilon\to 0}\begin{pmatrix}\epsilon^{-1}\boldsymbol{I} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{I}\end{pmatrix} \qquad (5.3.16)\\
&\qquad\times[\boldsymbol{I}-(\boldsymbol{I}\otimes\boldsymbol{\kappa})\{\boldsymbol{\alpha}^{*c}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}^{c}+\boldsymbol{\alpha}^{*c}\boldsymbol{U}^{*c}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}^{c}\epsilon+\boldsymbol{\alpha}^{*c}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}^{c}\boldsymbol{D}^{c}\epsilon+o(\epsilon)\}]\\
&\quad=-\begin{pmatrix}(\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{U}^{*c}_{\sigma u}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}^{c}_{u\sigma}+\boldsymbol{D}^{c}_{\sigma\sigma} & (\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{U}^{*c}_{\sigma+}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}^{c}_{+-}+\boldsymbol{D}^{c}_{\sigma-}\\ (\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}_{-u}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}^{c}_{u\sigma} & (\boldsymbol{I}\otimes\boldsymbol{\kappa})\boldsymbol{\alpha}^{*c}_{-+}(\boldsymbol{I}\otimes\boldsymbol{e})\boldsymbol{\beta}^{c}_{+-}-\boldsymbol{I}\end{pmatrix}. \qquad (5.3.17)
\end{aligned}
$$

Notice that (5.3.17) is the negative of a sub–intensity matrix, so in particular is nonsingular. □

Remark 17 *The approach used to compute the limits in the proof of Theorem 5.21 is somewhat similar to a L'Hospital–type of rule for matrices. However, in our case "differentation" is only performed in certain rows of the matrices, the ones corresponding to the states in which the Brownian component is active.*

5.4 Erlangization

In this section we explain how to use the method of erlangization to approximate the probability of Parisian and cumulative Parisian ruin with deterministic clocks.

As pointed out in Ramaswami et al. (2008), Jagerman (1982) provides a method to approximate any C^4 function $f : \mathbb{R}_+ \to \mathbb{R}_+$ as a limit of a sequence of functions $\{f_n\}_{n\geq 1}$ on the form

$$f_n(a) = \int_0^\infty \frac{(n/a)^n}{(n-1)!} s^{n-1} e^{-ns/a} f(s)\mathrm{d}s, \quad a > 0, n \geq 1. \tag{5.4.1}$$

In fact, the convergence $f_n \to f$ is uniform over compacts. Moreover,

$$f_n(a) - f(a) = \frac{1/n}{2} f^{(2)}(a) + O(1/n^2). \tag{5.4.2}$$

Furthemore, notice that

$$f_n(a) = \mathbb{E}(f(L^{(n)})), \tag{5.4.3}$$

where $L^{(n)}$ is a an Erlang distributed r.v. with n stages and mean a. This approximation method is known as **erlangization**. In the context of risk theory, it was first proposed in Asmussen et al. (2002), where given 5.4.2, the authors proposed to use the Richardson extrapolation sequence $\{\widehat{f}_n\}_{n\geq 1}$ with

$$\widehat{f}_n = (n+1)f_{n+1} - nf_n$$

to approximate f; (see Appendix C). The sequence $\{\widehat{f}_n\}_{n\geq 1}$ is such that the pointwise error of approximation to f is on the order $O(1/n^2)$, an order of speed higher than that of $\{f_n\}_{n\geq 1}$. Let us show how to use the erlangization method to approximate the probability of Parisian and cumulative Parisian ruin with deterministic clock(s).

Let $\{R_t\}_{t\geq 0}$ be a fluid flow risk with $\mathfrak{E}_\sigma = \emptyset$. Define

$$g_i(u,a) = \phi_i^{\delta_a}(u), \quad a > 0, u \geq 0, i \in \mathfrak{E}_\sigma \cup \mathfrak{E}_+,$$

so that $g_i(u,a)$ corresponds to the probability of cumulative Parisian ruin with deterministic clock of length a, and conditional on $R_0 = u, J_0 = i$. By conditioning on the first jump time of $\{J_t\}_{t\geq 0}$, it can be readily proved that the mapping $a \to g_i(u,a)$ is infinitely differentiable. This and (5.4.3) implies that $g_i(u,a)$ can be approximated by the sequence $\{\phi_i^{E^n}(u)\}_{n\geq 1}$ where E^n denotes the Erlang distribution function with n stages and mean a. We can take a step further and apply the Richardson extrapolation formula to improve the speed of convergence to $g_i(u,a)$. This is summarized in the following.

Theorem 5.22 *Let $\{R_t\}$ be a fluid flow risk process with $\mathfrak{E}_\sigma = \emptyset$ and fix $a > 0, u \geq 0$ and $i \in \mathfrak{E}_+$. Then we can compute the probability of cumulative Parisian ruin by*

$$\phi_i^{\delta_a}(u) = \lim_{n\to\infty} \phi_i^{E^n}(u) \tag{5.4.4}$$

$$= \lim_{n\to\infty} \widehat{\phi}_i^{E^n}(u), \tag{5.4.5}$$

where

$$\widehat{\phi}_i^{E^n}(u) = (n+1)\phi_i^{E^{n+1}}(u) - n\phi_i^{E^n}(u), \quad n \geq 1$$

and $\phi_i^{E^n}(u)$ is on the form (5.2.15) for the case $F = E^n$ corresponding to an Erlang distribution with n stages and mean a. Moreover, the convergence of (5.4.4) is on the order $O(1/n)$ and the the convergence of (5.4.5) is on the order $O(1/n^2)$.

The case of Parisian ruin is slightly more complex. The reason is that $\psi_i^F(u)$ cannot be thought as a deterministic function evaluated at a random point. Indeed, Parisian ruin has not one, but several i.i.d. clocks associated to each subexcursion below 0. However, in a similar fashion to that of (5.4.4), we get that for $\bar{\psi}_{ij}(F)$ as defined by

$$\bar{\psi}_{ij}(F) = \mathbb{P}(L > h, J_\tau = j \mid V_0 = 0, J_0 = i)$$

with $L \sim F, L \perp \{(V_t, J_t)\}_{t\geq 0}, i \in \mathfrak{E}_-, j \in \mathfrak{E}_+$, and $\bar{\boldsymbol{\Psi}}(F) := \{\bar{\psi}_{ij}(F)\}_{\mathfrak{E}_-, j\in\mathfrak{E}_+}$, we have that

$$\bar{\boldsymbol{\Psi}}(\delta_a) = \bar{\boldsymbol{\Psi}}(E^n) + \frac{\boldsymbol{C}}{n} + O(1/n^2),$$

where E^n corresponds to an Erlang distribution with n stages and mean a, and $\boldsymbol{C}$ is some real matrix. In the following we analyse how this is related to the probability of Parisian ruin.

Theorem 5.23 *Let $\{R_t\}$ be a fluid flow risk process with $\mathfrak{E}_\sigma = \emptyset$ and fix $a > 0, u \geq 0$ and $i \in \mathfrak{E}_+$. Then we can compute the probability of Parisian ruin by*

$$\psi_i^{\delta_a}(u) = \lim_{n\to\infty} \psi_i^{E^n}(u) \tag{5.4.6}$$

$$= \lim_{n\to\infty} \widehat{\psi}_i^{E^n}(u), \tag{5.4.7}$$

where

$$\widehat{\psi}_i^{E^n}(u) = (n+1)\psi_i^{E^{n+1}}(u) - n\psi_i^{E^n}(u), \quad n \geq 1$$

and $\psi_i^{E^n}(u)$ is on the form (5.2.13) for the case $F = E^n$ corresponding to an Erlang distribution with n stages and mean a. Moreover, the convergence of (5.4.6) is on the order $O(1/n)$ and the the convergence of (5.4.7) is on the order $O(1/n^2)$.

PROOF. From the matrix formula

$$(\boldsymbol{A}_1 - \boldsymbol{A}_2)^{-1} = \boldsymbol{A}_1^{-1} + \boldsymbol{A}_1^{-1}(\boldsymbol{I} - \boldsymbol{A}_2\boldsymbol{A}_1^{-1})^{-1}\boldsymbol{A}_2\boldsymbol{A}_1^{-1},$$

by substituting with $\boldsymbol{A}_1 = \boldsymbol{I} - \bar{\boldsymbol{\Psi}}(\delta_a)$ and $\boldsymbol{A}_2 = \frac{\boldsymbol{C}\boldsymbol{\beta}_{+-}}{n} + O(1/n^2)$, we get that

$$\begin{aligned}\left(\boldsymbol{I} - \bar{\boldsymbol{\Psi}}\left(E^n\right)\boldsymbol{\beta}_{+-}^c\right)^{-1} &= \left(\boldsymbol{I} - \bar{\boldsymbol{\Psi}}(\delta_a) - \frac{\boldsymbol{C}\boldsymbol{\beta}_{+-}}{n} + O(1/n^2)\right)^{-1} \\ &= \boldsymbol{A}_1^{-1} + \boldsymbol{A}_1^{-1}(\boldsymbol{I} - \boldsymbol{A}_4\boldsymbol{A}_1^{-1})^{-1}\boldsymbol{A}_4\boldsymbol{A}_1^{-1} \\ &= \boldsymbol{A}_1^{-1} + \boldsymbol{A}_1^{-1}(\boldsymbol{I} + O(1/n))^{-1}\left(\frac{\boldsymbol{C}\boldsymbol{\beta}_{+-}}{n} + O(1/n^2)\right)\boldsymbol{A}_1^{-1}.\end{aligned}$$

Since $\boldsymbol{I} + O(1/n) \to \boldsymbol{I}$ as $n \to \infty$, then there exists $n_0 \geq 1$ such that for all $n \geq n_0$, the term $(\boldsymbol{I} + O(1/n))^{-1}$ exists and its operator norm is bounded, say, by 2. This means that there exists some matrix $\widehat{\boldsymbol{C}}$ such that

$$\begin{aligned}\left(\boldsymbol{I} - \bar{\boldsymbol{\Psi}}\left(E^n\right)\boldsymbol{\beta}_{+-}^c\right)^{-1} &= \boldsymbol{A}_1^{-1} + \frac{\boldsymbol{D}}{n} + O(1/n^2) \\ &= \left(\boldsymbol{I} - \bar{\boldsymbol{\Psi}}\left(\delta_a\right)\boldsymbol{\beta}_{+-}^c\right)^{-1} + \frac{\widehat{\boldsymbol{C}}}{n} + O(1/n^2)\end{aligned}$$

Using Theorem 5.5 we get that

$$\begin{aligned}\psi_i^{E^n}(u) &= \boldsymbol{e}_i'\boldsymbol{\beta}_{+-}^c e^{\boldsymbol{D}^c u}\left(\left(\boldsymbol{I} - \bar{\boldsymbol{\Psi}}\left(\delta_a\right)\boldsymbol{\beta}_{+-}^c\right)^{-1} + \frac{\widehat{\boldsymbol{C}}}{n} + O(1/n^2)\right) \\ &\quad \times \left(\boldsymbol{e} - \left(\bar{\boldsymbol{\Psi}}(\delta_a) - \frac{\boldsymbol{C}}{n} + O(1/n^2)\right)\boldsymbol{e}\right) \\ &= \psi_i^{\delta_a}(u) + \frac{\gamma}{n} + O(1/n^2),\end{aligned}$$

for some $\gamma \in \mathbb{R}$. This proves (5.4.6). (5.4.7) follows by performing a Richardson approximation (see Appendix C). □

Remark 18 *Theorem 5.22 and Theorem 5.23 propose to use an Erlang approximation to the probability of cumulative Parisian and Parisian ruin with deterministic clocks. Is this the best approach possible, considering the tools provided by Corollary 5.7 and Theorem 5.9? The answer is affirmative and is backed up by the fact that within the class of phase–type distributions of dimension n, E^n is the least variable; see Aldous and Shepp (1987) and O'Cinneide (1991b).*

We refer to Bladt et al. (2018) for a numerical assessment of the erlangization method to approximate the probability of Parisian and cumulative Parisian ruin with deterministic clocks.

5.5 Conclusions and remarks

In this chapter we provided a method to compute the Parisian and cumulative Parisian probability of ruin for a fairly general class of risk processes. Below is a brief analysis of the methods existing in the literature and how they are compared to ours.

- Parisian ruin has been extensively studied before, mainly for spectrally negative Lévy process. For instance, Parisian ruin with deterministic clocks was first studied in Dassios and Wu (2008) for the Cramér–Lundberg process. Their results were later generalized to the case of general spectrally negative Lévy processes in Czarna and Palmowski (2011) and Loeffen et al. (2013) for the case of deterministic clocks, and in Landriault et al. (2013) and Baurdoux et al. (2016) for the case of random clocks, Erlang and exponential, respectively. In Czarna et al. (2017), the distribution of the number of claims leading to Parisian ruin is computed for the Cramér-–Lundberg process. A related source is Albrecher and Ivanovs (2017), where a discretely observed ruin model is declared ruined if it is ever below 0 at its observation points driven by a Poisson process; this coincides with the case of Parisian ruin with exponentially distributed clocks. In comparison, cumulative Parisian ruin with a deterministic clock has had far less attention, with some resouces being Landriault et al. (2011) and Guérin and Renaud (2017). The majority of these papers use the theory of scale functions; see Kuznetsov et al. (2012) for a comprehensive study. The methodology presented in this chapter is completely different from theirs.
- Our approach relies heavily in the matrix–analytic nature of the fluid flow risk process and phase–type distributions. In particular, our approach is inspired by the erlangization method proposed in Asmussen et al. (2002) to study finite–time probability of ruin for a Cramér–Lundberg process, which later got generalized in Stanford et al. (2005) and Ramaswami et al. (2008). To the best of our knowledge, this is the first attempt to study Parisian and cumulative Parisian ruin within a matrix–analytic framework.
- Our method provides a completely explicit formula to compute the probability of Parisian and cumulative parisian ruin for phase–type–distributed

clocks, as well as an approximation method to compute the case of deterministic clocks. In the Parisian setting, an erlangization method was previously proposed by Landriault et al. (2013) for the case of spectrally negative Lévy processes. Although some of our models overlap with models of Landriault et al. (2013) (for instance, spectrally negative Lévy processes with phase–type claims), in general we study different models and different perspectives. Moreover, ours is the first instance in which the use of the Richardson approximation is proposed to improve the speed of convergence of the erlangization, within the Parisian or cumulative Parisian setting.

Chapter 6

Some excursion properties of spectrally negative Lévy processes over a matrix–exponential horizon

6.1 Introduction

Risk theory has been considerably benefited by the advances on fluctuation theory for Lévy processes. One these advances is the so–called Wiener–Hopf decomposition, which corresponds to a distributional property of the excursions from the maximum of the Lévy process prior to an independent exponential random time. In particular, if the Lévy process happens to have negative jumps only, the Wiener–Hopf factorisation has a special form in terms of analytic functions.

The aim of this chapter is to extend the Wiener–Hopf factorisation in order to study the distributional properties of the excursions from the maximum of a spectrally negative Lévy process prior to an independent matrix–exponential random time. In order to do so, we use a technique based on holomorphic func-

tional calculus for matrices, which broadly speaking, lets us "evaluate analytic functions at a matrix". Later on, we give an additional example of how holomorphic functional calculus for matrices can be applied to the study of cumulative Parisian ruin for spectrally negative Lévy processes. We finalise this chapter with a discussion of our results with previous ones existing in the literature.

The main original results stemming from this chapter are Theorem 6.6 and Corollary 6.10.

6.2 Definition and properties of the holomorphic functional calculus for matrices

Functional calculus corresponds to the theory which allows to apply mathematical functions to linear operators. Several types of functional calculi can be developed for different kinds of linear operators acting on different kind of spaces. In this section we specialize in the holomorphic functional calculus for bounded linear operators in Banach spaces developed in Dunford and Schwartz (1958), which states that for any analytic function $f : U \subset \mathbb{C} \to \mathbb{C}$ and linear operator $\boldsymbol{A}$ with spectrum $\mathrm{sp}\{A\} \in U$, the linear operator $f(\boldsymbol{A})$ is defined by

$$f(\boldsymbol{A}) = \frac{1}{2\pi \mathrm{i}} \oint_{\gamma} f(z)(z\boldsymbol{I} - \boldsymbol{A})^{-1} \mathrm{d}z,$$

where $\gamma \subset U$ is a clock–wise oriented closed chain which encloses $\mathrm{sp}(\boldsymbol{A})$. From now on, we will restrict ourselves to linear operators acting on $\mathbb{C}^n$, so that the linear operators are simply complex–valued square matrices. A few basic properties of this functional calculus are the following. We include their respective proof in Appendix B.

Theorem 6.1 *Let f, g be two analytic functions on $B \subset \mathbb{C}$ and let $\boldsymbol{A}$ be a complex–valued matrix such that $\mathrm{sp}(\boldsymbol{A}) \subset B$. Then:*

1. $f(\boldsymbol{A}) + g(\boldsymbol{A}) = (f + g)(\boldsymbol{A})$.
2. $f(\boldsymbol{A})g(\boldsymbol{A}) = (fg)(\boldsymbol{A})$.
3. $\mathrm{sp}(f(\boldsymbol{A})) = f(\mathrm{sp}(\boldsymbol{A}))$.
4. *If* $\mathrm{sp}(g(\boldsymbol{A})) \subset B$, *then* $f(g(\boldsymbol{A})) = (f \circ g)(\boldsymbol{A})$.

Remark 19 *Holomorphic functional calculus coincides with the traditional definitions of functions of matrices. For example, for any nonsingular matrix $\boldsymbol{A}$*

and $n \in \mathbb{Z}$, $\boldsymbol{A}^n$ coincides with the matrix $f(\boldsymbol{A})$ with $f(z) = z^n$. Also, the matrix exponential of $\boldsymbol{A}$ defined as

$$e^{\boldsymbol{A}} = \sum_{i=0}^{\infty} \frac{\boldsymbol{A}^i}{i!}$$

coincides with the matrix $g(\boldsymbol{A})$ with $g(z) = e^z$. This and more details on the subject can be found in Doolittle (1998).

6.3 On an extension of the Wiener–Hopf factorisation for spectrally negative Lévy processes

Let $\{X_t\}_{t\geq 0}$ be a spectrally negative Lévy process with $X_0 = 0$ and Lévy exponent $\psi(\theta) := \log \mathbb{E}(e^{\theta} X_1)$ given by

$$\psi(\theta) = a\theta + \frac{1}{2}\sigma^2\theta^2 + \int_{(-\infty,0)} (e^{\theta x} - 1 - \theta x \mathbb{1}_{x>-1})\Pi(\mathrm{d}x). \tag{6.3.1}$$

Let $\tau_x^+ = \inf\{s > 0 : X_s > x\}$. We call $\{\tau_x^+\}_{x\geq 0}$ the **first–passage time process**. The following is a summary of the properties of the process $\{\tau_x^+\}_{x\geq 0}$ (see Kyprianou (2014), Theorem 3.12 and Theorem 3.14).

Theorem 6.2 *If $\{\tau_x^+\}_{x\geq 0}$ is the first–passage time process of a spectrally negative Lévyprocess $\{X_t\}_{t\geq 0}$ with Lévy exponent $\psi(\cdot)$, then:*

- *For all $q \geq 0$,*
$$\mathbb{E}(e^{-q\tau_x^+}) = e^{-\phi(q)x},$$
where $\phi(q) = \sup\{a \geq 0 : \psi(a) = q\}$, that is, ϕ corresponds to the right–inverse of ψ.
- *The process $\{\tau_x^+\}_{x\geq 0}$ is a non–decreasing Lévy process with Lévy exponent given by ϕ.*
- *For all $q > 0$, the function ϕ can be expressed as*
$$\phi(q) = a_\tau q + \int_{(0,\infty)} (1 - e^{-qx})\Pi_\tau(\mathrm{d}x).z \tag{6.3.2}$$

Theorem 6.3 *The function ϕ can be analytically extended to $D := \{z \in \mathbb{C} : \mathrm{Re}(z) > 0\}$. Furthermore, for any $z_0 \in D$, $\mathrm{Re}(\phi(z_0)) > 0$.*

PROOF. Extend the definition (6.3.2) of ϕ to D. To verify that ϕ is analytic in D, let γ be a closed curve in D. Then,

$$\begin{aligned}\oint_\gamma \phi(z)\mathrm{d}z &= a_\tau \oint_\gamma z\mathrm{d}z + \oint_\gamma \int_{(0,\infty)} (1-e^{-zx})\Pi_\tau(\mathrm{d}x)\mathrm{d}z \\ &= 0 + \int_{(0,\infty)} \left(\oint_\gamma (1-e^{-zx})\mathrm{d}z\right)\Pi_\tau(\mathrm{d}x) \\ &= \int_{(0,\infty)} 0\cdot\Pi_\tau(\mathrm{d}x) = 0,\end{aligned}$$

so that Morera's theorem imply that ϕ is indeed analytic in D. Additionally, for $z_0 = a_0 + ib_0 \in D$ ($a_0 \in (0,\infty), b_0 \in \mathbb{R}$),

$$\mathrm{Re}(\phi(z_0)) = a_\tau a_0 + \int_{(0,\infty)} (1-\cos(b_0 x)e^{-a_0 x})\Pi_\tau(\mathrm{d}x),$$

and since $1-\cos(b_0x)e^{-a_0x} > 0$ for all $x > 0$, then $\mathrm{Re}(\phi(z_0)) > 0$. Next we present a result which helps us understand how the process $\{X_t\}_{t\geq 0}$ and its maximum behaves prior to an independent exponential random time; such a result is known as the **Wiener–Hopf factorisation**. The term "Wiener–Hopf" is not restricted to the study of Lévy processes. In fact, it was first proposed as a technique to solve a convolution equation in Wiener and Paley (1934). From there on, intepretations and applications of this technique were used in diverse areas such as in complex analysis, differential equations and operator theory. In probability theory, the Wiener–Hopf techique was used in Spitzer (1957) to study maxima of a random walk, which was later studied in Rogozin (1966) for Lévy processes. The following corresponds to the Wiener–Hopf factorisation for spectrally negative Lévy process (see Subsection 6.5.2 in Kyprianou (2014)).

Theorem 6.4 *Fix $p > 0$. Let $T \sim \mathrm{Exp}(p)$ be independent of the spectrally negative Lévy process $\{X_t\}_{t\geq 0}$ with Lévy exponent $\psi(\cdot)$, and for $t \geq 0$ define*

$$\overline{X}_t = \max_{0\leq s\leq t} X_s, \qquad \overline{G}_t = \sup\{s < t : X_s = \overline{X}_s\}.$$

Then,

$$(\overline{G}_T, \overline{X}_T) \perp (T - \overline{G}_T, X_T - \overline{X}_T).$$

Furthermore, there exists $\delta = \delta(p) > 0$ such that for all $\nu, \theta \in [0,\delta)$

$$p - \nu > 0, \quad \phi(p-\nu) - \theta > 0, \quad p - \nu - \psi(\theta) > 0, \tag{6.3.3}$$

$$\begin{aligned}\mathbb{E}(e^{\nu\overline{G}_T + \theta\overline{X}_T}) &= \frac{\phi(p)}{\phi(p-\nu)-\theta}, \quad \textit{and} \\ \mathbb{E}(e^{\nu(T-\overline{G}_T)+\theta(X_T-\overline{X}_T)}) &= \frac{p}{\phi(p)}\frac{\phi(p-\nu)-\theta}{p-\nu-\psi(\theta)},\end{aligned}$$

where $\phi(\cdot)$ is defined as in Theorem 6.2, In particular,

$$\mathbb{E}(e^{\nu T+\theta X_T}) = \frac{p}{p-\nu-\psi(\theta)}.$$

The goal of this section is to study $\mathbb{E}(e^{\nu\overline{G}_Y+\theta\overline{X}_Y})$, $\mathbb{E}(e^{\nu(Y-\overline{G}_Y)+\theta(X_Y-\overline{X}_Y)})$ and $\mathbb{E}(e^{\nu Y+\theta X_Y})$, where Y is a matrix–exponential random variable independent from $\{X_t\}_{t\geq 0}$ and suitable $\nu,\theta\geq 0$. After we do this, we will be able to figure out if $(\overline{G}_T,\overline{X}_T)\perp(T-\overline{G}_T,X_T-\overline{X}_T)$ or if any kind of factorisation holds. The way we proceed is by using holomorphic functional calculus. The following corresponds to Theorem 4.2 of Bladt et al. (2015).

Theorem 6.5 *Let $Y\sim \mathrm{ME}(\boldsymbol{\pi},\boldsymbol{T},\boldsymbol{t})$ and let $w:\mathbb{R}\to\mathbb{R}$ be a function with Laplace transform*

$$L_w(s) := \int_0^\infty e^{-sx}w(x)\mathrm{d}x$$

which exists and is finite for at least some $s_0\in[0,-\mathrm{Re}(dev(\boldsymbol{T})))$. Furthermore, define

$$L_w^*(s) := sL_w(s) = \int_0^\infty se^{-sx}w(x)\mathrm{d}x.$$

Then, $\mathbb{E}(w(Y))$ can be computed in two ways:

$$\begin{aligned}\mathbb{E}(w(Y)) &= \boldsymbol{\pi}L_w(-\boldsymbol{T})\boldsymbol{t}, \qquad \textit{or}\\ \mathbb{E}(w(Y)) &= -\boldsymbol{\pi}L_w^*(-\boldsymbol{T})\boldsymbol{e}.\end{aligned}$$

Proof. Since $L_w(s_0)$ exists and is finite, then it also exists and is finite for all $s>s_0$. Define $D_{s_0}=s_0+D=\{z\in\mathbb{C}:\mathrm{Re}(z)>s_0\}$. Then L_w can be analytically extended to D_{s_0}. Indeed, if $\gamma_0\subset D_{s_0}$ is a closed curve, then

$$\begin{aligned}\int_{\gamma_0}L_w(z)\mathrm{d}z &= \int_{\gamma_0}\left(\int_0^\infty e^{-zx}w(x)\mathrm{d}x\right)\mathrm{d}z\\ &= \int_0^\infty\left(\int_{\gamma_0}e^{-zx}\mathrm{d}z\right)w(x)\mathrm{d}x\\ &= \int_0^\infty 0\cdot w(x)\mathrm{d}x = 0,\end{aligned}$$

so that by Morera's theorem, L_w is analytic in D_{s_0}. Furthermore, $\mathrm{sp}(-\boldsymbol{T})\subset$

D_{s_0}. Then, if $\gamma \subset D_{s_0}$ is a closed chain which encloses $\mathrm{sp}(-\boldsymbol{T})$,

$$\begin{aligned}
\mathbb{E}(w(Y)) &= \int_0^\infty w(x)(\boldsymbol{\pi} e^{\boldsymbol{T}x}\boldsymbol{t})\mathrm{d}x \\
&= \boldsymbol{\pi}\left(\int_0^\infty w(x)e^{\boldsymbol{T}x}\mathrm{d}x\right)\boldsymbol{t} \\
&= \boldsymbol{\pi}\left(\int_0^\infty w(x)e^{-(-\boldsymbol{T})x}\mathrm{d}x\right)\boldsymbol{t} \\
&= \boldsymbol{\pi}\left(\int_0^\infty w(x)\left[\int_\gamma e^{-zx}(z\boldsymbol{I}-(-\boldsymbol{T}))^{-1}\mathrm{d}z\right]\mathrm{d}x\right)\boldsymbol{t} \\
&= \boldsymbol{\pi}\left[\int_\gamma \left(\int_0^\infty w(x)e^{-zx}\mathrm{d}x\right)(z\boldsymbol{I}-(-\boldsymbol{T}))^{-1}\mathrm{d}z\right]\boldsymbol{t} \\
&= \boldsymbol{\pi}\left[\int_\gamma L_w(z)(z\boldsymbol{I}-(-\boldsymbol{T}))^{-1}\mathrm{d}z\right]\boldsymbol{t} \\
&= \boldsymbol{\pi}L_w(-\boldsymbol{T})\boldsymbol{t}.
\end{aligned}$$

Now, $L_w^*(z) = zL_w(z)$ is also analytic in D_{s_0}, so that,

$$\begin{aligned}
\mathbb{E}(w(Y)) &= \boldsymbol{\pi}\left[\int_\gamma \left(\int_0^\infty w(x)e^{-zx}\mathrm{d}x\right)(z\boldsymbol{I}-(-\boldsymbol{T}))^{-1}\mathrm{d}z\right]\boldsymbol{t} \\
&= \boldsymbol{\pi}\left[\int_\gamma z^{-1}\left(\int_0^\infty zw(x)e^{-zx}\mathrm{d}x\right)(z\boldsymbol{I}-(-\boldsymbol{T}))^{-1}\mathrm{d}z\right]\boldsymbol{t} \\
&= \boldsymbol{\pi}\left[\int_\gamma z^{-1}L_w^*(z)(z\boldsymbol{I}-(-\boldsymbol{T}))^{-1}\mathrm{d}z\right]\boldsymbol{t} \\
&= \boldsymbol{\pi}\boldsymbol{T}^{-1}L_w^*(-\boldsymbol{T})\boldsymbol{t} \\
&= \boldsymbol{\pi}L_w^*(-\boldsymbol{T})(\boldsymbol{T}^{-1})(-\boldsymbol{T}\boldsymbol{e}) \\
&= -\boldsymbol{\pi}L_w^*(-\boldsymbol{T})\boldsymbol{e},
\end{aligned}$$

and the proof is finished. □

The following is the factorisation correspondent to the process $\{X_t\}_{t\geq 0}$ prior to an independent matrix–exponential random time.

Theorem 6.6 *Let $Y \sim \mathrm{ME}(\boldsymbol{\pi}, \boldsymbol{T})$ be independent of the spectrally negative Lévy process $\{X_t\}_{t\geq 0}$. Then, there exists $\delta > 0$ such that for any $\nu, \theta \in [0, \delta)$,*

$$\mathrm{Re}(\mathrm{sp}(-\boldsymbol{T} - \nu\boldsymbol{I})) \subset (0, \infty), \tag{6.3.4}$$

$$\mathrm{Re}(\mathrm{sp}(\phi(-\boldsymbol{T} - \nu\boldsymbol{I}) - \theta\boldsymbol{I})) \subset (0, \infty), \tag{6.3.5}$$

$$\mathrm{Re}(\mathrm{sp}(-\boldsymbol{T} - (\nu + \psi(\theta))\boldsymbol{I})) \subset (0, \infty), \tag{6.3.6}$$

and for $\nu_1, \nu_2, \theta_1, \theta_2 \in [0, \delta)$,

$$\begin{aligned}&\mathbb{E}(\exp([\nu_1 \overline{G}_Y + \theta_1 \overline{X}_Y] + [\nu_2(Y - \overline{G}_Y) + \theta_2(X_Y - \overline{X}_Y]) \\ &\quad = -\boldsymbol{\pi} \boldsymbol{F}_1(\nu_1, \theta_1) \boldsymbol{F}_2(\nu_2, \theta_2) \boldsymbol{e}. \end{aligned} \tag{6.3.7}$$

where

$$\begin{aligned}&\boldsymbol{F}_1(\nu_1, \theta_1) = \phi(-\boldsymbol{T})(\phi(-\boldsymbol{T} - \nu_1 \boldsymbol{I}) - \theta_1 \boldsymbol{I})^{-1}, \quad \textit{and} \\ &\boldsymbol{F}_2(\nu_2, \theta_2) = (-\boldsymbol{T})(\phi(-\boldsymbol{T}))^{-1}(\phi(-\boldsymbol{T} - \nu_2 \boldsymbol{I}) - \theta_2 \boldsymbol{I})(-\boldsymbol{T} - (\nu_2 + \psi(\theta_2))\boldsymbol{I})^{-1}.\end{aligned}$$

In particular, for all $\nu, \theta \in [0, \delta)$,

$$\mathbb{E}(e^{\nu Y + \theta X_Y}) = -\boldsymbol{\pi}(-\boldsymbol{T})(-\boldsymbol{T} - (\nu + \psi(\theta))\boldsymbol{I}). \tag{6.3.8}$$

Proof.

Fix $\lambda_0 \in (0, -\mathrm{Re}(\mathrm{dev}(\boldsymbol{T})))$ and let δ be $\delta(\lambda_0)$ as in Theorem 6.4: this guarantees that (6.3.4), (6.3.5) and (6.3.6) hold. Fix $\nu_1, \nu_2, \theta_1, \theta_2 \in [0, \delta)$. If we define

$$w(t) := \mathbb{E}(\exp([\nu_1 \overline{G}_t + \theta_1 \overline{X}_t] + [\nu_2(t - \overline{G}_t) + \theta_2(X_t - \overline{X}_t]), \quad t \geq 0.$$

then $L_w^*(\lambda_0)$ is finite. Moreover, if $s \geq \lambda_0$ and $Z \sim \mathrm{Exp}(s)$ independent of $\{X_t\}_{t \geq 0}$,

$$\begin{aligned}L_w^*(s) &= \int_0^\infty s e^{-st} w(t) \mathrm{d}t \\ &= \mathbb{E}\left(w(Z)\right) \\ &= \mathbb{E}(\exp([\nu_1 \overline{G}_Z + \theta_1 \overline{X}_Z] + [\nu_2(Z - \overline{G}_Z) + \theta_2(X_Z - \overline{X}_Z)]) \\ &= \mathbb{E}(\exp(\nu_1 \overline{G}_Z + \theta_1 \overline{X}_Z)) \cdot \mathbb{E}(\exp(\nu_2(Z - \overline{G}_Z) + \theta_2(X_Z - \overline{X}_Z))) \\ &= \left[\frac{\phi(s)}{\phi(s - \nu_1) - \theta_1}\right] \left[\frac{s}{\phi(s)} \frac{\phi(s - \nu_2) - \theta_2}{s - \nu_2 - \psi(\theta_2)}\right]\end{aligned}$$

Using Theorem 6.5,

$$\begin{aligned}&\mathbb{E}(\exp([\nu_1 \overline{G}_Y + \theta_1 \overline{X}_Y] + [\nu_2(Y - \overline{G}_Y) + \theta_2(X_Y - \overline{X}_Y]) \\ &\quad = \mathbb{E}(w(Y)) = -\boldsymbol{\pi} L_w^*(-\boldsymbol{T}) \boldsymbol{e}. \end{aligned} \tag{6.3.9}$$

Apply Theorem 6.1 to finalise the proof. □

Remark 20 *Equation (6.3.7) implies that, in general,* $(\overline{G}_Y, \overline{X}_Y)$ *is* **not** *independent from* $(Y - \overline{G}_Y, X_Y - \overline{X}_Y)$*: this follows by noting that*

$$-\boldsymbol{\pi} \boldsymbol{F}_1(\nu_1, \theta_1) \boldsymbol{F}_2(\nu_2, \theta_2) \boldsymbol{e} \neq (-\boldsymbol{\pi} \boldsymbol{F}_1(\nu_1, \theta_1) \boldsymbol{e})(-\boldsymbol{\pi} \boldsymbol{F}_2(\nu_2, \theta_2) \boldsymbol{e}).$$

The following provides a way to construct nontrivial representations of matrix–exponential distributions based on the findings of Theorem 6.6.

Corollary 6.7 *Let ψ be the Lévy exponent of a spectrally negative Lévy process, and let ϕ be its right–inverse. Let $(\boldsymbol{\pi}, \boldsymbol{T})$ correspond to a representation of a matrix–exponential distribution. Then $(\boldsymbol{\pi}, -\phi(-\boldsymbol{T}))$ also corresponds to a representation of a matrix–exponential distribution.*

Proof. Let $\{X_t\}_{t\geq 0}$ be the Lévy process with Lévy exponent ψ, and let $Y \sim \mathrm{ME}(\boldsymbol{\pi}, \boldsymbol{T})$ be independent of $\{X_t\}_{t\geq 0}$. Replacing $\nu_1 = \nu_2 = \theta_2 = 0$ in (6.3.7) we get that

$$\begin{aligned}\mathbb{E}(e^{\theta_1 \overline{X}_Y}) &= -\boldsymbol{\pi}[\phi(-\boldsymbol{T})(\phi(-\boldsymbol{T}) - \theta_1 \boldsymbol{I})^{-1}]\boldsymbol{e} \\ &= \boldsymbol{\pi}[(\phi(-\boldsymbol{T}) - \theta_1 \boldsymbol{I})^{-1}](-\phi(-\boldsymbol{T})\boldsymbol{e}),\end{aligned}$$

which by analytic continuation implies that

$$L_{\overline{X}_Y}(z) := \mathbb{E}(e^{-z\overline{X}_Y}) = \boldsymbol{\pi}(z\boldsymbol{I} - (-\phi(-\boldsymbol{T})))^{-1}(-\phi(-\boldsymbol{T})\boldsymbol{e}), \quad z \in D.$$

Using the inverse Laplace transform formula, we get that the density of $f_{\overline{X}_Y}$ of $\overline{X}_Y$ is given by

$$\begin{aligned}f_{\overline{X}_Y}(x) &= \frac{1}{2\pi \mathrm{i}} \int_\gamma e^{xz} L_{\overline{X}_Y}(z) \mathrm{d}z \\ &= \boldsymbol{\pi} \left(\frac{1}{2\pi \mathrm{i}} \int_\gamma e^{xz} (z\boldsymbol{I} - (-\phi(-\boldsymbol{T})))^{-1} \mathrm{d}z \right) (-\phi(-\boldsymbol{T})\boldsymbol{e}) \\ &= \boldsymbol{\pi} e^{-\phi(-\boldsymbol{T})x} (-\phi(-\boldsymbol{T})\boldsymbol{e}), \quad x \geq 0,\end{aligned}$$

where γ denotes a closed curve contained in D which encloses $\mathrm{sp}(-\phi(-\boldsymbol{T}))$. This implies that $\overline{X}_Y \sim \mathrm{ME}(\boldsymbol{\pi}, -\phi(-\boldsymbol{T}))$ and the proof is finished. □

6.4 Cumulative Parisian ruin for spectrally negative Lévy processes

Let $\{X_t\}_{t\geq 0}$ be a spectrally negative Lévy process with $X_0 = 0$ and with Lévy exponent ψ, and let ϕ be the right–inverse of ψ. Let $Y \sim \mathrm{ME}(\boldsymbol{\pi}, \boldsymbol{T})$ be independent of $\{X_t\}_{t\geq 0}$. In Chapter 5 we defined cumulative Parisian ruin with clock Y as the event $\{\int_0^\infty \mathbb{1}\{X_s \leq 0\}\mathrm{d}s > Y\}$. In other words, cumulative Parisian ruin happens if the occupation time of $\{X_t\}_{t\geq 0}$ below 0 is larger than Y; this

is nontrivial only in the case $\{X_t\}_{t\geq 0}$ drifts to $+\infty$, which is what we assume from now on. The following is a result from Landriault et al. (2011) regarding the Laplace transform of the total occupation time of a spectrally negative spectrally negative Lévy process.

Theorem 6.8 *If $\{X_t\}_{t\geq 0}$ is a spectrally negative Lévy process with Lévy exponent ψ with right–inverse ϕ. Then, for $\lambda \geq 0$,*

$$\mathbb{E}\left(\exp\left(-\lambda \int_0^\infty \mathbb{1}\{X_s \leq 0\}\mathrm{d}s\right)\right) = \psi'(0+)\frac{\phi(\lambda)}{\lambda}. \tag{6.4.1}$$

The previous provides a characterisation of the distribution of $\int_0^\infty \mathbb{1}\{X_s \leq 0\}\mathrm{d}s$. In the following we prove how, in general, one can relate the Laplace transform of a random variable to an independent matrix–exponential–distributed random variable using holomorphic functional calculus.

Theorem 6.9 *Let W be any nonnegative random variable with Laplace transform $L(s) = \mathbb{E}(e^{-sW})$ and let $Y \sim \mathrm{ME}(\boldsymbol{\pi}, \boldsymbol{T})$ be independent of W. Then,*

$$\mathbb{P}(W \leq Y) = -\boldsymbol{\pi} L(-\boldsymbol{T})\boldsymbol{e}. \tag{6.4.2}$$

Proof. Define

$$w(x) := \mathbb{P}\left(W \leq x\right).$$

For fixed $s > 0$, let $Z \sim \mathrm{Exp}(s)$ be independent of W. Then,

$$\begin{aligned} L_w^*(s) &= \int_0^\infty s e^{-sx} w(x)\mathrm{d}x \\ &= \int_0^\infty s e^{-sx} \mathbb{P}\left(W \leq x\right)\mathrm{d}x \\ &= \mathbb{P}\left(W \leq Z\right) \\ &= \mathbb{E}\left(\mathbb{P}\left(W \leq Z \mid W\right)\right) \\ &= \mathbb{E}\left(e^{-sW}\right) \\ &= L(s) \end{aligned}$$

Thus,

$$\begin{aligned} \mathbb{P}\left(W \leq Y\right) &= \mathbb{E}(w(Y)) \\ &= -\boldsymbol{\pi} L_w^*(-\boldsymbol{T})\boldsymbol{e} = -\boldsymbol{\pi} L(-\boldsymbol{T})\boldsymbol{e}, \end{aligned}$$

and 6.4.2 follows. □

The following provides a way to compute the probability of cumulative Parisian ruin with a matrix–exponential–distributed clock.

Corollary 6.10 *Let $\{X_t\}_{t\geq 0}$ be a spectrally negative Lévy process with Lévy exponent ψ with right–inverse ϕ. Let $Y \sim PH(\boldsymbol{\pi}, \boldsymbol{T})$ be independent of $\{X_t\}_{t\geq 0}$. Then,*

$$\mathbb{P}\left(\int_0^\infty \mathbb{1}\{X_s \leq 0\}\mathrm{d}s > \boldsymbol{Y}\right) = 1 + \psi'(0^+)\boldsymbol{\pi}(-\boldsymbol{T})^{-1}\phi(-\boldsymbol{T})\boldsymbol{e}. \tag{6.4.3}$$

Proof. Apply Theorem 6.9 with $W = \int_0^\infty \mathbb{1}\{X_s \leq 0\}\mathrm{d}s$ and then use Theorem 6.8. □

Remark 21 *If one desires to perform a Laplace inversion of (6.4.1), we can compute (6.4.3) with $Y \sim E^n$ where E^n denotes an Erlang distribution of n stages and mean t, and the limit as $n \to \infty$ would approach $\mathbb{P}(\int_0^\infty \mathbb{1}\{X_s \leq 0\}\mathrm{d}s > t)$. This would be no different from the inversion method of Jagerman (1982), which has been shown to converge to its limit at a rate of $O(1/n)$. We could go even further and apply a Richardson extrapolation to speed up the rate of convergence to $O(1/n^2)$ (see Appendix C). However, we have an advantage in Corollary 6.10: it is valid for* **all** *matrix–exponential distributions, not only for phase–type–distributed ones. In O'Cinneide (1991b) it is shown that the Erlang distribution of n stages is no longer the least variable amongst the class of matrix–exponential distributions of dimension n ($n \geq 3$). In Horváth et al. (2016) it is conjectured, through numerical experimentation, that the least variable matrix–exponential distribution of order n has a coefficient of variation similar to $2/n^2$. In comparison, the Erlang distribution of order n has coefficient of variation $1/n$. This would mean that, in theory, we would be able to approximate $\mathbb{P}(\int_0^\infty \mathbb{1}\{X_s \leq 0\}\mathrm{d}s > t)$ faster if we choose the sequence of least variable matrix–exponential distributions of order n instead of Erlang distributions of n stages; we could then apply Richardson extrapolation to that sequence in order to speed the convergence even further.*

6.5 Examples

In Theorem 6.6 and Corollary 6.10, one of the main ingredients is the matrix $\phi(\boldsymbol{S}_\alpha)$ where $\boldsymbol{S}_\alpha = -\boldsymbol{T} - \alpha\boldsymbol{I}$ for some $\alpha \in \mathbb{R}$ such that $\mathrm{Re}(\mathrm{sp}(\boldsymbol{S}_\alpha)) \subset (0, \infty)$. In this section we explicitly compute $\phi(\boldsymbol{S}_\alpha)$ for two examples: whenever $\boldsymbol{T}$ is the subintensity matrix associated to an Erlang distribution of n stages, and for $\boldsymbol{T}$ corresponding to a matrix–exponential distribution which is strictly non–phase–type.

6.5.1 Erlang clock

Let $Y \sim \mathrm{Erl}(n, \beta)$. In this case, $\boldsymbol{S}_\alpha$ is a $n \times n$ matrix on the form

$$\boldsymbol{S}_\alpha = -\boldsymbol{T} - \alpha\boldsymbol{I} = \begin{pmatrix} \beta-\alpha & -\beta & 0 & \cdots & 0 & 0 \\ 0 & \beta-\alpha & -\beta & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \beta-\alpha & -\beta \\ 0 & 0 & 0 & \cdots & 0 & \beta-\alpha \end{pmatrix}.$$

Notice that $\sigma(\boldsymbol{S}_\alpha) = \{\beta - \alpha\}$. Let $\gamma \subset D$ be a closed curve enclosing $\beta - \alpha$. It is straightforward to verify that $(z\boldsymbol{I} - \boldsymbol{S}_\alpha)^{-1}$ is equal to

$$\begin{pmatrix} \frac{1}{z-(\beta-\alpha)} & \frac{-\beta}{(z-(\beta-\alpha))^2} & \frac{\beta^2}{(z-(\beta-\alpha))^3} & \cdots & \frac{(-\beta)^{n-2}}{(z-(\beta-\alpha))^{n-1}} & \frac{(-\beta)^{n-1}}{(z-(\beta-\alpha))^n} \\ 0 & \frac{1}{z-(\beta-\alpha)} & \frac{-\beta}{(z-(\beta-\alpha))^2} & \cdots & \frac{(-\beta)^{n-3}}{(z-(\beta-\alpha))^{n-2}} & \frac{(-\beta)^{n-2}}{(z-(\beta-\alpha))^{n-1}} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{z-(\beta-\alpha)} & \frac{-\beta}{(z-(\beta-\alpha))^2} \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{z-(\beta-\alpha)} \end{pmatrix}.$$

Thus, in order to compute $\phi(\boldsymbol{S}_\alpha) = \frac{1}{2\pi\mathrm{i}} \oint_\gamma \phi(z)(z\boldsymbol{I} - \boldsymbol{S}_\alpha)^{-1}\mathrm{d}z$ it is enough to calculate (for all $i \in \{1, \ldots, n\}$)

$$(-\beta)^{k-1}\frac{1}{2\pi\mathrm{i}}\oint_\gamma \frac{\phi(z)}{(z-(\beta-\alpha))^k}\mathrm{d}z = (k-1)!(-\beta)^{k-1}\phi^{(k-1)}(\beta-\alpha), \quad (6.5.1)$$

where the Residue theorem is used in the last equality, with $\phi^{(k)}$ denoting the k-th derivative of h.

6.5.2 ME–distributed clock which is non–PH

Consider the density function

$$f(x) = 2e^{-x}(1 - \cos(x)).$$

Such a function corresponds to the $ME(\boldsymbol{\pi}, \boldsymbol{T})$ distribution with

$$\boldsymbol{\pi} = (1, 0, 0), \quad \boldsymbol{T} = \begin{pmatrix} -1 & -1 & 2/3 \\ 1 & -1 & -2/3 \\ 0 & 0 & -1 \end{pmatrix},$$

so that

$$\boldsymbol{S}_\alpha = -\boldsymbol{T} - \alpha\boldsymbol{I} = \begin{pmatrix} 1-\alpha & 1 & -2/3 \\ -1 & 1-\alpha & 2/3 \\ 0 & 0 & 1-\alpha \end{pmatrix}.$$

In this case, we have that $\mathrm{sp}(\boldsymbol{S}_\alpha) = \{1-\alpha, 1+\mathrm{i}-\alpha, 1-\mathrm{i}-\alpha\}$. Let $\gamma \subset D$ be a closed curve enclosing $\mathrm{sp}(\boldsymbol{S}_\alpha)$. Then, one can verify that $(z\boldsymbol{I} - \boldsymbol{S}_\alpha)^{-1}$ is equal to

$$\begin{pmatrix} \frac{z-(1-\alpha)}{(z-(1-\mathrm{i}-\alpha))(z-(1+\mathrm{i}-\alpha))} & \frac{1}{(z-(1-\mathrm{i}-\alpha))(z-(1+\mathrm{i}-\alpha))} & -\frac{2(z-(2-\alpha))}{3(z-(1-\alpha))(z-(1-\mathrm{i}-\alpha))(z-(1+\mathrm{i}-\alpha))} \\ -\frac{1}{(z-(1-\mathrm{i}-\alpha))(z-(1+\mathrm{i}-\alpha))} & \frac{z-(1-\alpha)}{(z-(1-\mathrm{i}-\alpha))(z-(1+\mathrm{i}-\alpha))} & \frac{2(z+\alpha)}{3(z-(1-\alpha))(z-(1-\mathrm{i}-\alpha))(z-(1+\mathrm{i}-\alpha))} \\ 0 & 0 & \frac{1}{z-(1-\alpha)} \end{pmatrix}.$$

By the Residue theorem, we have that

$$\phi(\boldsymbol{S}_\alpha) = \frac{1}{2\pi\mathrm{i}} \oint_\gamma \phi(z)(z\boldsymbol{I} - \boldsymbol{S}_\alpha)^{-1}\mathrm{d}z = \begin{pmatrix} a & b & -\frac{2}{3}c \\ -b & a & \frac{2}{3}d \\ 0 & 0 & e \end{pmatrix}, \tag{6.5.2}$$

where

$$\begin{aligned} a &= \frac{(1-\mathrm{i}-\alpha)-(1-\alpha)}{(1-\mathrm{i}-\alpha)-(1+\mathrm{i}-\alpha)}\phi(1-\mathrm{i}-\alpha) + \frac{(1+\mathrm{i}-\alpha)-(1-\alpha)}{(1+\mathrm{i}-\alpha)-(1-\mathrm{i}-\alpha)}\phi(1+\mathrm{i}-\alpha) \\ &= \frac{1}{2}\phi(1-\mathrm{i}-\alpha) + \frac{1}{2}\phi(1+\mathrm{i}-\alpha), \end{aligned}$$

$$\begin{aligned} b &= \frac{1}{(1-\mathrm{i}-\alpha)-(1+\mathrm{i}-\alpha)}\phi(1-\mathrm{i}-\alpha) + \frac{1}{(1+\mathrm{i}-\alpha)-(1-\mathrm{i}-\alpha)}\phi(1+\mathrm{i}-\alpha) \\ &= \frac{\mathrm{i}}{2}\phi(1-\mathrm{i}-\alpha) - \frac{\mathrm{i}}{2}\phi(1+\mathrm{i}-\alpha) \end{aligned}$$

$$\begin{aligned} c &= \frac{(1-\alpha)-(2-\alpha)}{((1-\alpha)-(1-\mathrm{i}-\alpha))((1-\alpha)-(1+\mathrm{i}-\alpha))}\phi(1-\alpha) \\ &\quad + \frac{(1-\mathrm{i}-\alpha)-(2-\alpha)}{((1-i-\alpha)-(1-\alpha))((1-\mathrm{i}-\alpha)-(1+\mathrm{i}-\alpha))}\phi(1-\mathrm{i}-\alpha) \\ &\quad + \frac{(1+\mathrm{i}-\alpha)-(2-\alpha)}{((1+\mathrm{i}-\alpha)-(1-\alpha))((1+\mathrm{i}-\alpha)-(1-\mathrm{i}-\alpha))}\phi(1+i-\alpha) \\ &= -\phi(1-\alpha) - \frac{-1-\mathrm{i}}{2}\phi(1-\mathrm{i}-\alpha) - \frac{-1+\mathrm{i}}{2}\phi(1+\mathrm{i}-\alpha) \end{aligned}$$

$$\begin{aligned} d &= \frac{(1-\alpha)+\alpha}{((1-\alpha)-(1-i-\alpha))((1-\alpha)-(1+\mathrm{i}-\alpha))}\phi(1-\alpha) \\ &\quad + \frac{(1-\mathrm{i}-\alpha)+\alpha}{((1-\mathrm{i}-\alpha)-(1-\alpha))((1-\mathrm{i}-\alpha)-(1+\mathrm{i}-\alpha))}\phi(1-\mathrm{i}-\alpha) \\ &\quad + \frac{(1+\mathrm{i}-\alpha)+\alpha}{((1+\mathrm{i}-\alpha)-(1-\alpha))((1+\mathrm{i}-\alpha)-(1-\mathrm{i}-\alpha))}\phi(1+i-\alpha) \\ &= \phi(1-\alpha) - \frac{1-\mathrm{i}}{2}\phi(1-\mathrm{i}-\alpha) - \frac{1+\mathrm{i}}{2}\phi(1+\mathrm{i}-\alpha), \quad \text{and} \end{aligned}$$

$$e = \phi(1-\alpha).$$

Equation (6.5.2) can also be written in compact form as

$$\begin{aligned}\phi(\boldsymbol{S}_\alpha) &= \frac{1}{2\pi \mathrm{i}} \oint_\gamma \phi(z)(z\boldsymbol{I} - \boldsymbol{S}_\alpha)^{-1}\mathrm{d}z \\ &= \begin{pmatrix} 0 & 0 & 2/3 \\ 0 & 0 & 2/3 \\ 0 & 0 & 1 \end{pmatrix} \phi(1-\alpha) + \begin{pmatrix} 1/2 & \mathrm{i}/2 & (-1-\mathrm{i})/3 \\ -\mathrm{i}/2 & 1/2 & (-1+\mathrm{i})/3 \\ 0 & 0 & 0 \end{pmatrix} \phi(1-\mathrm{i}-\alpha) \\ &\quad + \begin{pmatrix} 1/2 & -\mathrm{i}/2 & (-1+\mathrm{i})/3 \\ \mathrm{i}/2 & 1/2 & (-1-\mathrm{i})/3 \\ 0 & 0 & 0 \end{pmatrix} \phi(1+\mathrm{i}-\alpha).\end{aligned}$$

6.6 Conclusions and remarks

In this chapter we showed how holomorphic functional calculus can be used to revisit results from fluctuation theory of spectrally negative Lévy processes. Below is a brief analysis of our findings and their comparison with the existing literature.

- The use of functional calculus in relation to matrix–exponential distributions was first proposed in Asmussen and Perry (1998), where a functional calculus based on matrix power series was used to study a reflected Brownian motion stopped at an matrix–exponential random time. In Bladt et al. (2015), the authors used holomorphic functional calculus to implement a general method to compute means of real functions, like $\sin(x)$ or x^α ($\alpha \in (0,1)$), evaluated at a matrix–exponential–distributed random point. In Subsection 8.5.3 of Bladt and Nielsen (2017), the authors study a Brownian motion inspected at a matrix–exponential–distributed random time using holomorphic functional calculus. The approach we followed regarding functional calculus is very much in line with the one of Bladt et al. (2015) and Subsection 8.5.3 of Bladt and Nielsen (2017).
- The Wiener–Hopf factorisation of Lévy processes over a phase–type horizon was recently studied in Asmussen and Ivanovs (2018), where the authors show that a conditional Wiener–Hopf factorisation holds using time reversal of the underlying Markov jump process. For our study of the Wiener–Hopf factorisation over a matrix–exponential horizon, we do not use time reversal nor conditioning in order to arrive to the desired result, but rather a purely abstract method to compute the formulae. On one hand, we do provide a completely explicit factorisation in a straightforward way; the downside is that we are not able to provide a deep probabilistic insight of the process.

- The study of cumulative Parisian ruin in this chapter differs from the one in Chapter 5 in two things. Firstly, here we study a "different" kind class of risk process: spectrally negative Lévy processes in some cases fall under the category of fluid flow risk processes but not in general. Moreover, in the framework of this chapter we allow the clock to be matrix–exponential–distributed: this required the use of abstract theory rather than probabilistic argumentations.

Chapter 7

Fluid RAPs

7.1 Introduction

In the present chapter we consider an extension of the fluid flow process without Brownian components. Inspired by the physical interpretation of the RAP in Subsection 2.3.2, we study a Markov additive process, which we coin Fluid RAP (or FRAP), whose additive component is still piecewise linear, but its background component is now an orbit process. The main contribution is showing that some results of the fluid flow process presented in Section 2.4 translate verbatim into the FRAP framework, although new methods and techniques need to be used in doing so.

The structure of this chapter is as follows. First, we define the Proto–FRAP, a process which is simple in its nature, however, it will help us introduce the physical interpretation of the orbit process and some of its restrictions. Then we give a precise definition of the FRAP, along with some key distributional properties. Later on, we compute first passage probabilities of the FRAP. We conclude this chapter by discussing some benefits of the model, and its similarities and diferences with the ones existing in the literature.

Unless otherwise stated, all the results in this chapter are original.

7.2 An example of a Fluid RAP

In this section we introduce the simplest non-trivial example of what we aim to model. Fix $m, n \geq 1$, let $\mathfrak{U}$ and $\mathfrak{D}$ be subsets of an affine subspace of $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively.[1] In some instances, we regard $\mathfrak{U} \cup \mathfrak{D}$ not as a geometrical object, but rather as a set union of $\mathfrak{U}$ and $\mathfrak{D}$, whose elements are row vectors of dimension either m or n. For any $a, b \geq 1$, let $M^{a\times b}(\mathbb{R})$ denote the family of $a \times b$-dimensional matrices with real entries. Let

$$\boldsymbol{\mathcal{T}} = \begin{pmatrix} \boldsymbol{C}_+ & \boldsymbol{D}_{+-} \\ \boldsymbol{C}_- & \boldsymbol{D}_{-+} \end{pmatrix},$$

with $\boldsymbol{C}_+ \in M^{m\times m}(\mathbb{R})$, $\boldsymbol{D}_{+-} \in M^{m\times n}(\mathbb{R})$, $\boldsymbol{D}_{-+} \in M^{n\times m}(\mathbb{R})$, and $\boldsymbol{C}_- \in M^{n\times n}(\mathbb{R})$ be such that $\boldsymbol{\mathcal{T}}\boldsymbol{e} = \boldsymbol{0}$.

Let $\{\boldsymbol{A}(t)\}_{t\geq 0}$ be a càdlàg piecewise deterministic Markov process with state–space $\mathfrak{U} \cup \mathfrak{D}$; that is, $\{\boldsymbol{A}(t)\}$ will be a row vector process of varying dimension, either m or n. The process $\{\boldsymbol{A}(t)\}_{t\geq 0}$ is such that $\boldsymbol{A}(0) = \boldsymbol{\alpha} \in \mathfrak{U} \cup \mathfrak{D}$, and for $t \geq 0$ in between jumps, it moves according to the system of differential equations given by

$$\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t} = \begin{cases} \boldsymbol{A}(t)\boldsymbol{C}_+ - \boldsymbol{A}(t)\boldsymbol{C}_+\boldsymbol{e}\cdot\boldsymbol{A}(t) & \text{for} \quad \boldsymbol{A}(t) \in \mathfrak{U}, \\ \boldsymbol{A}(t)\boldsymbol{C}_- - \boldsymbol{A}(t)\boldsymbol{C}_-\boldsymbol{e}\cdot\boldsymbol{A}(t) & \text{for} \quad \boldsymbol{A}(t) \in \mathfrak{D}. \end{cases} \tag{7.2.1}$$

As the process evolves in $\mathfrak{U} \cup \mathfrak{D}$, its jump intensity $\lambda : \mathfrak{U} \cup \mathfrak{D} \to \mathbb{R}_+$ is given by

$$\lambda(\boldsymbol{a}) = \begin{cases} \boldsymbol{a}\boldsymbol{D}_{+-}\boldsymbol{e} & \text{if} \quad \boldsymbol{a} \in \mathfrak{U}, \\ \boldsymbol{a}\boldsymbol{D}_{-+}\boldsymbol{e} & \text{if} \quad \boldsymbol{a} \in \mathfrak{D}, \end{cases} \tag{7.2.2}$$

and given that a jump occurs at time t, it will directly jump to

$$\frac{\boldsymbol{A}(t^-)\boldsymbol{D}_{+-}}{\boldsymbol{A}(t^-)\boldsymbol{D}_{+-}\boldsymbol{e}} \in \mathfrak{D} \quad \text{if} \quad \boldsymbol{A}(t^-) \in \mathfrak{U}, \quad \text{and}$$
$$\frac{\boldsymbol{A}(t^-)\boldsymbol{D}_{-+}}{\boldsymbol{A}(t^-)\boldsymbol{D}_{-+}\boldsymbol{e}} \in \mathfrak{U} \quad \text{if} \quad \boldsymbol{A}(t^-) \in \mathfrak{D}.$$

This means that $\{\boldsymbol{A}(t)\}_{t\geq 0}$ is càdlàg process which, in a piecewise sense, alternately takes values in $\mathfrak{U}$ and $\mathfrak{D}$. Such a process is inspired in the Rational arrival process (RAP) constructed in Subsection 2.3.2. Below are some properties inherited from RAPs.

Lemma 7.1 *Let $\{\boldsymbol{A}(t)\}_{t\geq 0}$ be a PDMP as described earlier.*

[1] If $m = n$, we regard $\mathfrak{U}$ and $\mathfrak{D}$ as subsets of different planes; this can be achieved by appropiately embedding them into two parallel n-dimensional planes in $\mathbb{R}^{n+1}$.

1. *On the event* $\{\boldsymbol{A}(\cdot)$ *has no jumps in* $[t, t+h]\}$,

$$\boldsymbol{A}(t+h) = \begin{cases} \frac{\boldsymbol{A}(t)e^{\boldsymbol{C}_+ h}}{\boldsymbol{A}(t)e^{\boldsymbol{C}_+ h}\boldsymbol{e}} & \text{if} \quad \boldsymbol{A}(t) \in \mathfrak{U}, \\ \frac{\boldsymbol{A}(t)e^{\boldsymbol{C}_- h}}{\boldsymbol{A}(t)e^{\boldsymbol{C}_- h}\boldsymbol{e}} & \text{if} \quad \boldsymbol{A}(t) \in \mathfrak{D}. \end{cases} \tag{7.2.3}$$

2.

$$\mathbb{P}(\boldsymbol{A}(\cdot) \text{ has no jumps in } [t, t+h] \mid \boldsymbol{A}(t)) = \begin{cases} \boldsymbol{A}(t)e^{\boldsymbol{C}_+ h}\boldsymbol{e} & \text{if} \quad \boldsymbol{A}(t) \in \mathfrak{U}, \\ \boldsymbol{A}(t)e^{\boldsymbol{C}_- h}\boldsymbol{e} & \text{if} \quad \boldsymbol{A}(t) \in \mathfrak{D}. \end{cases}$$

Proof.

1. Simply differentiate (7.2.3) w.r.t. h and verify that (7.2.1) holds.
2. Fix $t \geq 0$ and w.l.o.g. suppose $\boldsymbol{A}(t) \in \mathfrak{U}$. According to (2.1.3) and (7.2.3),

$$\mathbb{P}(\boldsymbol{A}(\cdot) \text{ has no jumps in } [t, t+h] \mid \boldsymbol{A}(t)) = F(h)$$

with

$$\begin{aligned} F(h) &= \exp\left(-\int_0^h \lambda\left(\frac{\boldsymbol{A}(t)e^{\boldsymbol{C}_+ h}}{\boldsymbol{A}(t)e^{\boldsymbol{C}_+ h}\boldsymbol{e}}\right)\right) \\ &= \exp\left(-\int_0^h \frac{\boldsymbol{A}(t)e^{\boldsymbol{C}_+ h}\boldsymbol{D}\boldsymbol{e}}{\boldsymbol{A}(t)e^{\boldsymbol{C}_+ h}\boldsymbol{e}}\right) \\ &= \exp\left(\int_0^h \frac{\boldsymbol{A}(t)e^{\boldsymbol{C}_+ h}\boldsymbol{C}\boldsymbol{e}}{\boldsymbol{A}(t)e^{\boldsymbol{C}_+ h}\boldsymbol{e}}\right), \end{aligned}$$

where the last equality followed because $\boldsymbol{D}\boldsymbol{e} = -\boldsymbol{C}\boldsymbol{e}$. Notice that

$$\frac{\mathrm{d}\ln(F(h))}{\mathrm{d}h} = \frac{\boldsymbol{A}(t)e^{\boldsymbol{C}_+ h}\boldsymbol{C}\boldsymbol{e}}{\boldsymbol{A}(t)e^{\boldsymbol{C}_+ h}\boldsymbol{e}} = \frac{\mathrm{d}\ln(\boldsymbol{A}(t)e^{\boldsymbol{C}_+ h}\boldsymbol{e})}{\mathrm{d}h}. \tag{7.2.4}$$

Since $F(0) = 1 = \boldsymbol{A}(t)e^{\boldsymbol{C}_+ 0}\boldsymbol{e}$, then (7.2.4) implies that $F(h) = \boldsymbol{A}(t)e^{\boldsymbol{C}_+ h}\boldsymbol{e}$ and the proof is finished.

Define the **Proto-FRAP** $\{(V_t, \boldsymbol{A}(t))\}_{t\geq 0}$ by

$$V_t = \int_0^t \mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{U}\} - \mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{D}\}\mathrm{d}s.$$

See Figure 7.1 for a visual description of the Proto-FRAP.

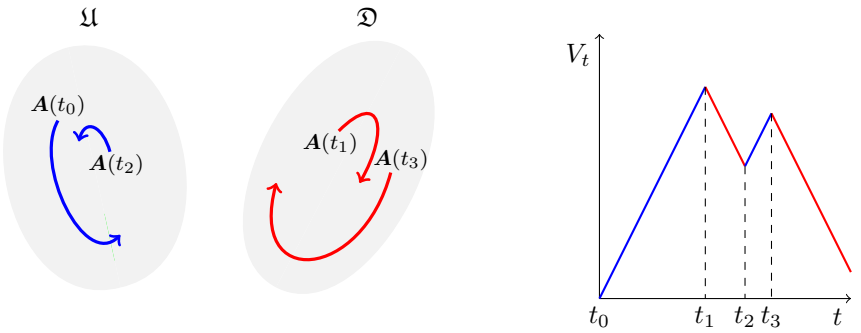


Figure 7.1: A sample path of a Proto-FRAP. $\{\boldsymbol{A}(t)\}_{t\geq 0}$ switches between states in $\mathfrak{U}$ and $\mathfrak{D}$, while $\{V_t\}_{t\geq 0}$ linearly increases or decreases accordingly.

Remark 22 *The following are some necessary conditions regarding the parameters and state–space of* $\{\boldsymbol{A}(t)\}_{t\geq 0}$.

1. $\boldsymbol{a}\boldsymbol{D}_{+-}\boldsymbol{e} \geq 0$ *for all* $\boldsymbol{a} \in \mathfrak{U}$ *and* $\boldsymbol{a}\boldsymbol{D}_{-+}\boldsymbol{e} \geq 0$ *for all* $\boldsymbol{a} \in \mathfrak{D}$. *This is to guarantee that the jump intensity* $\lambda(\cdot)$ *in (7.2.2) is well-defined.*
2. $\mathfrak{U}$ *and* $\mathfrak{D}$ *are bounded. In Subsection 2.3.2, this is a property needed for* $\{\boldsymbol{A}(t)\}_{t\geq 0}$ *in order to correspond to the coefficients of a linear combination of probability measures which is itself a probability measure. In our setting, it will enable us to use the Bounded Convergence Theorem in specific places and to guarantee that there is no infinite jump activity on compacts time intervals.*
3. *The process* $\{\boldsymbol{A}(t)\}_{t\geq 0}$ *is such that*

$$\lim_{t\to\infty} \mathbb{P}(\boldsymbol{A}(s) \notin \mathfrak{D} \text{ for all } s \leq t \mid \boldsymbol{A}(0) \in \mathfrak{U}) = 0, \quad \text{and}$$

$$\lim_{t\to\infty} \mathbb{P}(\boldsymbol{A}(s) \notin \mathfrak{U} \text{ for all } s \leq t \mid \boldsymbol{A}(0) \in \mathfrak{D}) = 0.$$

   This guarantees that $\{\boldsymbol{A}(t)\}_{t\geq 0}$ *is actually random and does jump between* $\mathfrak{U}$ *and* $\mathfrak{D}$.
4. $\boldsymbol{\mathcal{T}}$ *is such that* $\{\boldsymbol{A}(t)\}_{t\geq 0}$ *evolves within* $\mathfrak{U}\cup\mathfrak{D}$ *a.s. starting within any point in* $\mathfrak{U}\cup\mathfrak{D}$. *Although easy to state, this condition is not trivial to verify. For non–MAP or non–ME–renewal examples it needs to be verified on a case by case basis.*
5. $\mathfrak{U}$ *is contained in a minimal* $(m-1)$*–dimensional affine subspace of* $\mathbb{R}^m$ *and* $\mathfrak{D}$ *is contained in a minimal* $(n-1)$*–dimensional affine subspace of* $\mathbb{R}^n$. *We say that they are* **minimal** *in the sense that they cannot be contained in* $(m-2)$*– and* $(n-2)$*–dimensional affine subspaces, respectively. This is*

needed to assert that, whenever we have equations of the kind $\boldsymbol{aG} = \boldsymbol{aF}$ *for all* $\boldsymbol{a} \in \mathfrak{U}$*, it must happen that* $\boldsymbol{G} = \boldsymbol{F}$ *(ditto* $\boldsymbol{a} \in \mathfrak{D}$*). This is stated and proved precisely in Lemma 7.2 below.*

Lemma 7.2 *Let* $\boldsymbol{b} \in M^{m\times n}(\mathbb{R})$ *be such that*

$$\boldsymbol{ab} = \boldsymbol{0} \text{ for all } \boldsymbol{a} \in \mathfrak{U}. \tag{7.2.5}$$

Then $\boldsymbol{b} = \boldsymbol{0}$.

Proof. *Since* $\mathfrak{U}$ *is contained in a minimal* $(m-1)$*-dimensional affine space, then there must exist linearly independent vectors* $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{m-1} \in \mathbb{R}^m$ *and* $\boldsymbol{h} \in \mathfrak{U}$ *such that* $\{\boldsymbol{a}_i\}$ *are linearly independent from* $\boldsymbol{h}$ *and*

$$\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{m-1}\} + \boldsymbol{h} \subset \mathfrak{U} \subset \operatorname{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{m-1}\} + \boldsymbol{h}.$$

Thus, if we let

$$\boldsymbol{h}_1 = \boldsymbol{a}_1 + \boldsymbol{h}, \ldots, \boldsymbol{h}_{m-1} = \boldsymbol{a}_{m-1} + \boldsymbol{h}, \boldsymbol{h}_m = \boldsymbol{h},$$

then $\{\boldsymbol{h}_i\} \subset \mathfrak{U}$ *is a collection of linearly independent vectors that span* $\mathbb{R}^m$*. Let*

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{h}_1 \\ \vdots \\ \boldsymbol{h}_m \end{pmatrix}.$$

Then, (7.2.5) implies that

$$\boldsymbol{Hb} = \boldsymbol{0},$$

so that premultiplying the previous by $\boldsymbol{H}^{-1}$ *proves that* $\boldsymbol{b} = \boldsymbol{0}$. □

7.3 General model definition

Inspired by the example set in Section 7.2, in the following we define the general Markov additive process $\{(V_t, \boldsymbol{A}(t))\}_{t\geq 0}$ which we eventually coin Fluid RAP (FRAP). Let $\mathcal{P}$, $\mathcal{N}$ and $\mathcal{Z}$ be finite collections of ordered indices, and let $\mathcal{I} : \{+,-,0\} \to \{\mathcal{P},\mathcal{N},\mathcal{Z}\}$ be defined by $\mathcal{I}(+) = \mathcal{P}$, $\mathcal{I}(-) = \mathcal{N}$ and $\mathcal{I}(0) = \mathcal{Z}$. Also, let $\mathcal{S} = \cup_{k\in\{+,-,0\}}(\{k\} \times \mathcal{I}(k))$ and let $m(\cdot,\cdot)$ be some function $m : \mathcal{S} \to \{1, 2, \ldots\}$. For each $(k, i) \in \mathcal{S}$, let $\mathfrak{M}(k, i)$ be a subset of an affine subspace of $\mathbb{R}^{m(k,i)}$, and define

$$\begin{aligned} \mathfrak{U}_i &= \mathfrak{M}(+, i) \quad \text{for} \quad i \in \mathcal{P}, \\ \mathfrak{D}_i &= \mathfrak{M}(-, i) \quad \text{for} \quad i \in \mathcal{N}, \quad \text{and} \\ \mathfrak{Z}_i &= \mathfrak{M}(0, i) \quad \text{for} \quad i \in \mathcal{Z}. \end{aligned}$$

Furthermore, for $k \in \{+,-,0\}$ let $\mathfrak{M}(k) = \cup_{j\in\mathcal{I}(k)}\mathfrak{M}(k,j)$, and let $\mathfrak{U} = \mathfrak{M}(+)$, $\mathfrak{D} = \mathfrak{M}(-)$ and $\mathfrak{Z} = \mathfrak{M}(0)$.

For each $k \in \{+,-,0\}$ and $i,j \in \mathcal{I}(k)$, let $\boldsymbol{C}_{ij}^k \in M^{m(k,i)\times m(k,j)}(\mathbb{R})$. Additionally, for each $k,\ell \in \{+,-,0\}$ with $\ell \neq k$, $i \in \mathcal{I}(k)$ and $j \in \mathcal{I}(\ell)$, let $\boldsymbol{D}_{ij}^{k\ell} \in M^{m(k,i)\times m(\ell,j)}(\mathbb{R})$. Moreover, for each $k \in \{+,-,0\}$ let $d(k) = \#\{\mathcal{I}(k)\}$ correspond to the cardinality of $\mathcal{I}(k)$. Define

$$\boldsymbol{\mathcal{T}} = \begin{pmatrix} \boldsymbol{\mathcal{C}}_+ & \boldsymbol{\mathcal{D}}_{+-} & \boldsymbol{\mathcal{D}}_{+0} \\ \boldsymbol{\mathcal{D}}_{-+} & \boldsymbol{\mathcal{C}}_- & \boldsymbol{\mathcal{D}}_{-0} \\ \boldsymbol{\mathcal{D}}_{0+} & \boldsymbol{\mathcal{D}}_{0-} & \boldsymbol{\mathcal{C}}_0 \end{pmatrix},$$

where

$$\boldsymbol{\mathcal{C}}_k = \begin{pmatrix} \boldsymbol{C}_{11}^k & \boldsymbol{C}_{12}^k & \cdots & \boldsymbol{C}_{1,d(k)}^k \\ \boldsymbol{C}_{21}^k & \boldsymbol{C}_{22}^k & \cdots & \boldsymbol{C}_{2,d(k)}^k \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{C}_{d(k),1}^k & \boldsymbol{C}_{d(k),2}^k & \cdots & \boldsymbol{C}_{d(k),d(k)}^k \end{pmatrix}, \quad \text{and}$$

$$\boldsymbol{\mathcal{D}}_{k\ell} = \begin{pmatrix} \boldsymbol{D}_{11}^{k\ell} & \boldsymbol{D}_{12}^{k\ell} & \cdots & \boldsymbol{D}_{1,d(\ell)}^{k\ell} \\ \boldsymbol{D}_{21}^{k\ell} & \boldsymbol{D}_{22}^{k\ell} & \cdots & \boldsymbol{D}_{2,d(\ell)}^{k\ell} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{D}_{d(k),1}^{k\ell} & \boldsymbol{D}_{d(k),2}^{k\ell} & \cdots & \boldsymbol{D}_{d(k),d(\ell)}^{k\ell} \end{pmatrix}$$

for $k,\ell \in \{+,-,0\}$ with $\ell \neq k$. Moreover, assume that $\boldsymbol{\mathcal{T}}\boldsymbol{e} = \boldsymbol{0}$.

We define $\{\boldsymbol{A}(t)\}_{t\geq 0}$ as a càdlàg PDMP such that $\{\boldsymbol{A}(t)\}_{t\geq 0} \subset \mathfrak{U}\cup\mathfrak{D}\cup\mathfrak{Z}$, and for $t \geq 0$ between jumps, it moves according to the system of differential equations

$$\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t} = \boldsymbol{A}(t)\boldsymbol{C}_{ii}^k - \boldsymbol{A}(t)\boldsymbol{C}_{ii}^+\boldsymbol{e}\cdot\boldsymbol{A}(t) \quad \text{for} \quad k \in \{+,-,0\}, i \in \mathcal{I}(k), \boldsymbol{A}(t) \in \mathfrak{M}(k,i);$$

As in Lemma 7.1, this implies that on the event $\{\boldsymbol{A}(\cdot)$ has no jumps in $[t,t+h]$, $\boldsymbol{A}(t) \in \mathfrak{M}(k,i)\}$ with $k \in \{+,-,0\}, i \in \mathcal{I}(k)$,

$$\boldsymbol{A}(t+h) = \frac{\boldsymbol{A}(t)e^{\boldsymbol{C}_{ii}^k h}}{\boldsymbol{A}(t)e^{\boldsymbol{C}_{ii}^k h}\boldsymbol{e}}.$$

As the process $\{\boldsymbol{A}(t)\}_{t\geq 0}$ evolves, at each instant its jump intensity $\lambda : \mathfrak{U}\cup\mathfrak{D}\cup\mathfrak{Z} \to \mathbb{R}_+$ is given by

$$\lambda(\boldsymbol{a}) = -\boldsymbol{a}\boldsymbol{C}_{ii}^k\boldsymbol{e} = \sum_{j\in\mathcal{I}(k)\setminus\{i\}} \boldsymbol{a}\boldsymbol{C}_{ij}^k\boldsymbol{e} + \sum_{\ell\in\{+,-,0\}\setminus\{k\}}\sum_{j\in\mathcal{I}(\ell)} \boldsymbol{a}\boldsymbol{D}_{ij}^{k\ell}\boldsymbol{e} \quad \text{for} \quad \boldsymbol{a} \in \mathfrak{M}(k,i), \tag{7.3.1}$$

for $k \in \{+,-,0\}, i \in \mathcal{I}(k), \boldsymbol{a} \in \mathfrak{M}(k,i)$, and given that a jump occurs, say at some time $t > 0$ with $\boldsymbol{A}(t^-) \in \mathfrak{M}(k,i)$, then

1. It will land on $\frac{\boldsymbol{A}(t^-)\boldsymbol{C}_{ij}^k}{\boldsymbol{A}(t^-)\boldsymbol{C}_{ij}^k\boldsymbol{e}} \in \mathfrak{M}(k,j)$ $(j \in \mathcal{I}(k))$, with probability $\frac{\boldsymbol{A}(t^-)\boldsymbol{C}_{ij}^k\boldsymbol{e}}{-\boldsymbol{A}(t^-)\boldsymbol{C}_{ii}^k\boldsymbol{e}}$, or
2. It will land on $\frac{\boldsymbol{A}(t^-)\boldsymbol{D}_{ij}^{k\ell}}{\boldsymbol{A}(t^-)\boldsymbol{D}_{ij}^{k\ell}\boldsymbol{e}} \in \mathfrak{M}(\ell,j)$ $(\ell \in \{+,-,0\} \setminus \{k\},\ j \in \mathcal{I}(\ell))$, with probability $\frac{\boldsymbol{A}(t^-)\boldsymbol{D}_{ij}^{k\ell}\boldsymbol{e}}{-\boldsymbol{A}(t^-)\boldsymbol{C}_{ii}^k\boldsymbol{e}}$.

Just as in Lemma 7.1, (7.3.1) implies that

$$\mathbb{P}(\boldsymbol{A}(\cdot) \text{ has no jumps in } [t,t+h] \mid \boldsymbol{A}(t) \in \mathfrak{M}(k,i)) = \boldsymbol{A}(t)e^{\boldsymbol{C}_{ii}^k h}\boldsymbol{e}$$

for $k \in \{+,-,0\}, i \in \mathcal{I}(k)$. Now we are ready to define the Fluid RAP.

Definition 7.3 We define the **Fluid RAP (FRAP)** to be the Markov additive process $\{(V_t, \boldsymbol{A}(t))\}_{t\geq 0}$ with

$$V_t = \int_0^t \mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{U}\} - \mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{D}\}\mathrm{d}s.$$

See Figure 7.2 for a visual description.

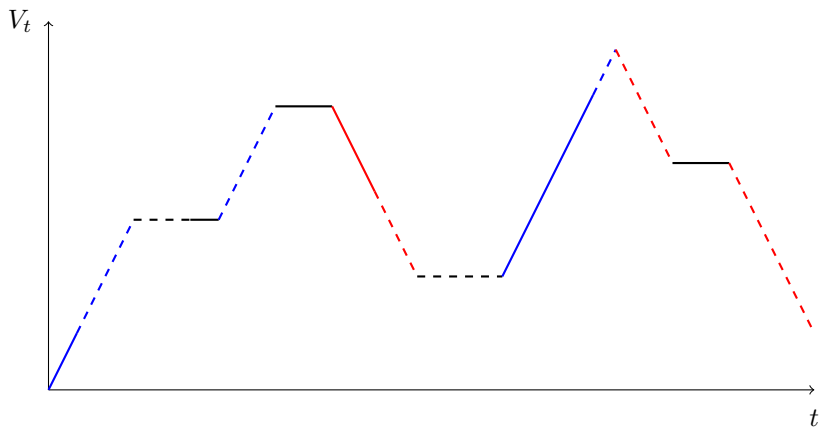


Figure 7.2: A sample path of a the process $\{V_t\}_{t\geq 0}$ whose associated process $\{\boldsymbol{A}(t)\}_{t\geq 0}$ has state–space $\mathfrak{U} \cup \mathfrak{D} \cup \mathfrak{Z}$ with $\mathfrak{U} = \mathfrak{U}_{\text{solid}} \cup \mathfrak{U}_{\text{dashed}}$, $\mathfrak{D} = \mathfrak{D}_{\text{solid}} \cup \mathfrak{D}_{\text{dashed}}$ and $\mathfrak{Z} = \mathfrak{Z}_{\text{solid}} \cup \mathfrak{Z}_{\text{dashed}}$.

Remark 23 *Just as in Section 7.2, we need the following conditions regarding the parameters and state–space of* $\{\boldsymbol{A}(t)\}_{t\geq 0}$.

1. $\boldsymbol{a}\boldsymbol{C}_{ij}^k\boldsymbol{e} \geq 0$ *for all* $(k,i) \in \mathcal{S}$, $j \in \mathcal{I}(\ell) \setminus \{i\}$, $\boldsymbol{a} \in \mathfrak{M}(k,i)$. *That is, the intensity of a jump within* $\mathfrak{U}$, $\mathfrak{D}$ *or* $\mathfrak{Z}$ *is also nonnegative. Furthermore,*

$\boldsymbol{a}\boldsymbol{D}_{ij}^{k\ell}\boldsymbol{e} \geq 0$ *for all* $(k,i) \in \mathcal{S}$, $\ell \neq k$, $j \in \mathcal{I}(\ell)$, $a \in \mathfrak{M}(k,i)$. *That is, the intensity of a jump happening between* $\mathfrak{U}$, $\mathfrak{D}$ *and* $\mathfrak{Z}$ *is nonnegative.*

2. $\{\mathfrak{U}_i\}_{i\in\mathcal{P}}$, $\{\mathfrak{D}_i\}_{i\in\mathcal{N}}$ *and* $\{\mathfrak{Z}_i\}_{i\in\mathcal{Z}}$ *are bounded.*
3. $\boldsymbol{\mathcal{T}}$ *is such that* $\{\boldsymbol{A}(t)\}_{t\geq 0}$ *evolves within* $\mathfrak{U}\cup\mathfrak{D}\cup\mathfrak{Z}$ *starting within any point in* $\mathfrak{U}\cup\mathfrak{D}\cup\mathfrak{Z}$.
4. *For* $k \in \{+,-\}$ *and* $\ell = \{k,0\}^c$,

$$\lim_{t\to\infty} \mathbb{P}(\boldsymbol{A}(s) \notin \mathfrak{M}(\ell) \text{ for all } s \leq t \mid \boldsymbol{A}(0) \in \mathfrak{M}(k) \cup \mathfrak{Z}) = 0.$$

5. *For all* $(k,i) \in \mathcal{S}$, $\mathfrak{M}(k,i)$ *is contained in a minimal* $(m(k,i)-1)$*-dimensional affine subspace of* $\mathbb{R}^{m(k,i)}$.

Unless otherwise stated, $\mathbb{P}$ ($\mathbb{E}$) denotes the probability (expectation) under the condition $\{V_0 = 0\}$. For $\boldsymbol{\alpha} \in \mathfrak{U}\cup\mathfrak{D}\cup\mathfrak{Z}$, we denote by $\mathbb{P}_{\boldsymbol{\alpha}}$ ($\mathbb{E}_{\boldsymbol{\alpha}}$) the probability (expectation) conditioned on the event $\{\boldsymbol{A}(0) = \boldsymbol{\alpha}, V_0 = 0\}$. Let $\Omega = \{\omega\} = \{(r,a)\} = \{\{(v_t,a_t)\}_{t\geq 0}\}$ be the canonical sample space of the paths of $\{(V_t,\boldsymbol{A}(t))\}_{t\geq 0}$. We call r the **level process** of ω and a the **orbit process** of ω. We say that $\omega = (r,a) \in \Omega$ is **shifted at level** $u \in \mathbb{R}$ whenever we refer to the path $\{(v_t+u,a_t)\}_{t\geq 0}$.

For the distributional analysis of the FRAP, it will be convenient to "encode" $\{\boldsymbol{A}(t)\}_{t\geq 0}$ in a larger (and sparser) row–vector process $\{\boldsymbol{B}(t)\}_{t\geq 0}$ which is defined as follows. First, for any row-vector $\boldsymbol{\beta}$ and $(k,i) \in \mathcal{S}$, define the function

$$b^{(k,i)}(\boldsymbol{\beta}) = \left(\mathbf{0}^{(k,1)}, \ldots, \mathbf{0}^{(k,i-1)}, \boldsymbol{\beta}, \mathbf{0}^{(k,i+1)}, \ldots, \mathbf{0}^{(k,d(k))}\right),$$

where $\mathbf{0}^{(k,j)}$ ($(k,j) \in \mathcal{S}$) is a row-vector of zeros with $m(k,j)$-elements. Then we define $\{\boldsymbol{B}(t)\}_{t\geq 0}$ by

$$\boldsymbol{B}(t) = \sum_{i\in\mathcal{I}(k)} b^{(k,i)}(\boldsymbol{A}(t)) \mathbb{1}\{\boldsymbol{A}(t) \in \mathfrak{M}(k,i)\} \quad \text{if } \boldsymbol{A}(t) \in \mathfrak{M}(k) \quad (t \geq 0).$$

Thus, the state–space of $\{\boldsymbol{B}(t)\}_{t\geq 0}$ will be $\cup_{k\in\{+,-,0\}}\mathfrak{M}^*(k)$ where

$$\mathfrak{M}^*(k) = \bigcup_{i=1}^{d(k)} \left(\prod_{j=1}^{i-1} \mathbf{0}^{(k,j)} \times \mathfrak{M}(k,i) \times \prod_{j=i+1}^{d(k)} \mathbf{0}^{(k,j)} \right).$$

The following result will be key to our analysis.

Theorem 7.4 *Let* $\boldsymbol{\Pi} : \mathbb{R}_+ \to M^{m\times m}(\mathbb{R})$ *be such that*

1. $\mathbf{\Pi}(\cdot)$ *is* $\|\cdot\|_{\max}$*–bounded in compact intervals of* $\mathbb{R}_+$*, and*
2. *For all* $x \geq 0$*,*

$$\mathbf{\Pi}(x) = e^{\boldsymbol{\mathcal{A}}x} + \int_0^x e^{\boldsymbol{\mathcal{A}}s}\boldsymbol{\mathcal{B}}\mathbf{\Pi}(x-s)\mathrm{d}s. \tag{7.3.2}$$

Then,

$$\mathbf{\Pi}(x) = e^{(\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})x} \quad \textit{for all} \quad x \geq 0.$$

Proof. Item 1 and the integral equation (7.3.2) imply that $\mathbf{\Pi}(\cdot)$ is infinitely differentiable. Thus, premultiplying (7.3.2) by $e^{-\boldsymbol{\mathcal{A}}x}$ and differentiating with respect to x gives us that

$$-e^{-\boldsymbol{\mathcal{A}}x}\boldsymbol{\mathcal{A}}\mathbf{\Pi}(x) + e^{-\boldsymbol{\mathcal{A}}x}\mathbf{\Pi}'(x) = e^{-\boldsymbol{\mathcal{A}}x}\boldsymbol{\mathcal{B}}\mathbf{\Pi}(x),$$

which is equivalent to

$$\mathbf{\Pi}'(x) = (\boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{B}})\mathbf{\Pi}(x).$$

Since we have the initial condition $\mathbf{\Pi}(0) = \boldsymbol{I}$, then $\mathbf{\Pi}(x) = e^{(\boldsymbol{\mathcal{A}}+\boldsymbol{\mathcal{B}})x}$ and the proof is finished.

Theorem 7.5 *For all* $k \in \{+,-,0\}$*, and on the event* $\boldsymbol{A}(0) \in \mathfrak{M}(k)$*,*

$$\mathbb{E}(\boldsymbol{B}(t)\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \textit{ for all } s \leq t\} \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)e^{\boldsymbol{\mathcal{C}}_k t} \quad (t \geq 0). \tag{7.3.3}$$

Proof. Let $\widehat{\boldsymbol{\mathcal{C}}}_k = \mathrm{diag}\{\boldsymbol{C}^k_{jj} : j \in \mathcal{I}(k)\}$. Define $t_0 = 0$ and for $n \geq 1$, let $t_n = \inf\{s > t_{n-1} : \boldsymbol{A}(s) \neq \boldsymbol{A}(s^-)\}$ be the epoch at which the n-th jump of $\{\boldsymbol{A}(t)\}_{t\geq 0}$ happens. For $t > 0$ let $L_t := \inf\{n \geq 0 : t_n < t\}$ be the number of jumps before time t. First, we show that for all $n \geq 0$ and $t > 0$,

$$\mathbb{E}(\boldsymbol{B}(t)\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \text{ for all } s \leq t, L_t = n\} \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)\mathbf{\Sigma}_n(t) \tag{7.3.4}$$

for unique continuous matrices $\{\mathbf{\Sigma}_n(t)\}_{n\geq 0}$. We do this by induction.

- **Case** $n = 0$**.**

$$\begin{aligned}
&\mathbb{E}(\boldsymbol{B}(t)\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \text{ for all } s \le t, L_t = 0\}) \\
&\quad = \mathbb{E}(\boldsymbol{B}(t) \mid \boldsymbol{A}(0), L_t = 0) \times \mathbb{P}(L_t = 0 \mid \boldsymbol{A}(0)) \\
&\quad = \sum_{i \in \mathcal{I}(k)} b^{(k,i)} \left(\frac{\boldsymbol{A}(0) e^{\boldsymbol{C}_{ii}^k t}}{\boldsymbol{A}(0) e^{\boldsymbol{C}_{ii}^k t} \boldsymbol{e}} \right) \times \boldsymbol{A}(0) e^{\boldsymbol{C}_{ii}^k t} \boldsymbol{e} \times \mathbb{1}\{\boldsymbol{A}(0) \in \mathfrak{M}(k,i)\} \\
&\quad = \sum_{i \in \mathcal{I}(k)} b^{(k,i)} (\boldsymbol{A}(0) e^{\boldsymbol{C}_{ii}^k t}) \times \mathbb{1}\{\boldsymbol{A}(0) \in \mathfrak{M}(k,i)\} \\
&\quad = \left[\sum_{i \in \mathcal{I}(k)} b^{(k,i)} (\boldsymbol{A}(0)) \times \mathbb{1}\{\boldsymbol{A}(0) \in \mathfrak{M}(k,i)\} \right] e^{\widehat{\boldsymbol{\mathcal{C}}}_k t} \\
&\quad = \boldsymbol{B}(0) e^{\widehat{\boldsymbol{\mathcal{C}}}_k t},
\end{aligned}$$

where the second–to–last equality followed by the block–diagonal structure of $\widehat{\boldsymbol{\mathcal{C}}}_k$. Thus, (7.3.4) is true if we choose $\boldsymbol{\Sigma}_0(t) = e^{\widehat{\boldsymbol{\mathcal{C}}}_k t}$, and this solution is unique by the minimality assumption of the FRAP; see Lemma 7.2.

- **Inductive part.** Suppose that (7.3.4) is true for some $n \ge 1$. Then

$$\begin{aligned}
&\mathbb{E}(\boldsymbol{B}(t)\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \text{ for all } s \le t, L_t = n+1\} \mid \boldsymbol{A}(0)) \\
&\quad = \int_0^t \mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(t)\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \text{ for all } s \le t, L_t = n+1, t_1 \in [r, r+\mathrm{d}r]\}) \\
&\quad = \int_0^t \mathbb{E}_{\boldsymbol{A}(0)} \left(\mathbb{E}_{\boldsymbol{A}(0)} \left(\boldsymbol{B}(t)\mathbb{1} \left\{ \begin{array}{c} \boldsymbol{A}(s) \in \mathfrak{M}(k) \text{ for all } s \le t, \\ L_t = n+1, t_1 \in [r, r+\mathrm{d}r] \end{array} \right\} \mid t_1, \boldsymbol{A}(t_1) \right) \right) \\
&\quad = \int_0^t \mathbb{E}_{\boldsymbol{A}(0)} \left(\begin{array}{c} \mathbb{E}_{\boldsymbol{A}(t_1)} (\boldsymbol{B}(t-r)\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \text{ for all } s \le t-r, L_{t-r} = n\}) \\ \times \mathbb{1}\{t_1 \in [r, r+\mathrm{d}r]\} \end{array} \right) \\
&\quad = \int_0^t \mathbb{E}_{\boldsymbol{A}(0)}([\boldsymbol{B}(t_1)\boldsymbol{\Sigma}_n(t-r)]\, \mathbb{1}\{t_1 \in [r, r+\mathrm{d}r]\}),
\end{aligned}$$

where the strong Markov property is used in the second–to–last equality, and the induction hypothesis in the last one. Now,

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(t_1)\mathbb{1}\{t_1 \in [0, \mathrm{d}r]\}) \\
&\quad = \sum_{i \in \mathcal{I}(k)} \sum_{j \in \mathcal{I}(k) \setminus \{i\}} b^{(k,j)} \left(\frac{\boldsymbol{A}(0)\boldsymbol{C}_{ij}^k}{\boldsymbol{A}(0)\boldsymbol{C}_{ij}^k \boldsymbol{e}} \right) (\boldsymbol{A}(0)\boldsymbol{C}_{ij}^k \boldsymbol{e}\mathrm{d}r)\mathbb{1}\{\boldsymbol{A}(0) \in \mathfrak{M}(k,i)\} \\
&\quad = \sum_{i \in \mathcal{I}(k)} \sum_{j \in \mathcal{I}(k) \setminus \{i\}} b^{(k,j)} \left(\boldsymbol{A}(0)\boldsymbol{C}_{ij}^k \right) \times \mathbb{1}\{\boldsymbol{A}(0) \in \mathfrak{M}(k,i)\}\mathrm{d}r \\
&\quad = \sum_{i \in \mathcal{I}(k)} b^{(k,i)} \left(\boldsymbol{A}(0) \right) \left(\boldsymbol{\mathcal{C}}_k - \widehat{\boldsymbol{\mathcal{C}}}_k \right) \times \mathbb{1}\{\boldsymbol{A}(0) \in \mathfrak{M}(k,i)\}\mathrm{d}r \\
&\quad = \boldsymbol{B}(0) \left(\boldsymbol{\mathcal{C}}_k - \widehat{\boldsymbol{\mathcal{C}}}_k \right) \mathrm{d}r,
\end{aligned}$$

so that

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(t_1)\mathbb{1}\{t_1 \in [r, r+\mathrm{d}r]\}) \\
&\quad = \mathbb{E}_{\boldsymbol{A}(0)}(\mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(t_1)\mathbb{1}\{t_1 \in [r, r+\mathrm{d}r]\} \mid \boldsymbol{A}(r)) \times \mathbb{1}\{L_r = 0\}) \\
&\quad = \mathbb{E}_{\boldsymbol{A}(0)}(\mathbb{E}_{\boldsymbol{A}(r)}(\boldsymbol{B}(t_1)\mathbb{1}\{t_1 \in [0, \mathrm{d}r]\}) \times \mathbb{1}\{L_r = 0\}) \\
&\quad = \mathbb{E}_{\boldsymbol{A}(0)}\left(\left[\boldsymbol{B}(r)\left(\boldsymbol{\mathcal{C}}_k - \widehat{\boldsymbol{\mathcal{C}}}_k\right)\right]\mathrm{d}r \times \mathbb{1}\{L_r = 0\}\right) \\
&\quad = \boldsymbol{B}(0)e^{\widehat{\boldsymbol{\mathcal{C}}}_k r}\left(\boldsymbol{\mathcal{C}}_k - \widehat{\boldsymbol{\mathcal{C}}}_k\right)\mathrm{d}r.
\end{aligned}$$

Thus,

$$\mathbb{E}(\boldsymbol{B}(t)\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \text{ for all } s \le t, L_t = n+1\} \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)\boldsymbol{\Sigma}_{n+1}(t)$$

with

$$\boldsymbol{\Sigma}_{n+1}(t) = \int_0^t e^{\widehat{\boldsymbol{\mathcal{C}}}_k r}\left(\boldsymbol{\mathcal{C}}_k - \widehat{\boldsymbol{\mathcal{C}}}_k\right)\boldsymbol{\Sigma}_n(t-r)\mathrm{d}r, \tag{7.3.5}$$

which happens to be continuous. Moreover, the uniqueness of $\boldsymbol{\Sigma}_{n+1}(\cdot)$ is guaranteed by the minimality of the FRAP.

Summing (7.3.5) over n and using Fubini's theorem (which is backed up by the boundedness of $\mathfrak{U} \cup \mathfrak{D} \cup \mathfrak{Z}$) we get that

$$\mathbb{E}_{\boldsymbol{\alpha}}(\boldsymbol{B}(t)\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \text{ for all } s \le t\}) = b^{(k,i)}(\boldsymbol{\alpha})\boldsymbol{\Sigma}(t),$$

where $\boldsymbol{\Sigma}(\cdot)$ is $\|\cdot\|_{\max}$–bounded and satisfies the equation

$$\boldsymbol{\Sigma}(t) = e^{\widehat{\boldsymbol{\mathcal{C}}}_k t} + \int_0^t e^{\widehat{\boldsymbol{\mathcal{C}}}_k s}\left(\boldsymbol{\mathcal{C}}_k - \widehat{\boldsymbol{\mathcal{C}}}_k\right)\boldsymbol{\Sigma}(t-s)\mathrm{d}s.$$

Thus, according to Theorem 7.4, $\boldsymbol{\Sigma}(t) = \exp\left(\left[\widehat{\boldsymbol{\mathcal{C}}}_k + (\boldsymbol{\mathcal{C}}_k - \widehat{\boldsymbol{\mathcal{C}}}_k)\right]t\right) = e^{\boldsymbol{\mathcal{C}}_k t}$ and (7.3.3) follows. □

Lemma 7.6 *For $k \in \{+,-,0\}$, let $\boldsymbol{A}(0) \in \mathfrak{I}(k)$. Define*

$$\rho_k = \inf\{s \ge 0 : \boldsymbol{A}(s) \notin \mathfrak{M}(k)\},$$

so that ρ_k is the **exit time from** *$\mathfrak{M}(k)$. Then, for $\ell \in \{+,-,0\} \setminus \{k\}$,*

1.

$$\mathbb{P}(\boldsymbol{A}(s) \in \mathfrak{M}(k) \text{ for all } s \le t \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)e^{\boldsymbol{\mathcal{C}}_k t}\boldsymbol{e}.$$

2.

$$\mathbb{E}(\boldsymbol{B}(\rho_k)\mathbb{1}\{\rho_k \in [t, t+\mathrm{d}t], \boldsymbol{A}(\rho_k) \in \mathfrak{M}(\ell)\} \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)e^{\boldsymbol{\mathcal{C}}_k t}\boldsymbol{\mathcal{D}}_{k\ell}\mathrm{d}t. \tag{7.3.6}$$

3.

$$\mathbb{E}(\boldsymbol{B}(\rho_k)\mathbb{1}\{\boldsymbol{A}(\rho_k) \in \mathfrak{M}(\ell)\} \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)(-\boldsymbol{\mathcal{C}}_k{}^{-1})\boldsymbol{\mathcal{D}}_{k\ell}. \tag{7.3.7}$$

Proof.

1. This is immediate from Theorem 7.5 and the fact that $\boldsymbol{B}(t)\boldsymbol{e} = 1$ for all $t \geq 0$.
2. First, notice that

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\rho_k)\mathbb{1}\{\rho_k \in [0, \mathrm{d}t], \boldsymbol{A}(\rho_k) \in \mathfrak{M}(\ell)\}) \\
&\quad = \sum_{i\in\mathcal{I}(k)}\sum_{j\in\mathcal{I}(\ell)} b^{(\ell,j)}\left(\frac{\boldsymbol{A}(0)\boldsymbol{D}^{k\ell}_{ij}}{\boldsymbol{A}(0)\boldsymbol{D}^{k\ell}_{ij}\boldsymbol{e}}\right) \times (\boldsymbol{A}(0)\boldsymbol{D}^{k\ell}_{ij}\boldsymbol{e}\mathrm{d}t) \times \mathbb{1}\{\boldsymbol{A}(0) \in \mathfrak{M}(k,i)\} \\
&\quad = \sum_{i\in\mathcal{I}(k)}\sum_{j\in\mathcal{I}(\ell)} b^{(\ell,j)}\left(\boldsymbol{A}(0)\boldsymbol{D}^{k\ell}_{ij}\right) \times \mathbb{1}\{\boldsymbol{A}(0) \in \mathfrak{M}(k,i)\}\mathrm{d}t \\
&\quad = \sum_{i\in\mathcal{I}(k)} b^{(k,i)}\left(\boldsymbol{A}(0)\right)\boldsymbol{\mathcal{D}}_{k\ell} \times \mathbb{1}\{\boldsymbol{A}(0) \in \mathfrak{M}(k,i)\}\mathrm{d}t \\
&\quad = \boldsymbol{B}(0)\boldsymbol{\mathcal{D}}_{k\ell}\mathrm{d}t,
\end{aligned}$$

so that

$$\begin{aligned}
&\mathbb{E}(\boldsymbol{B}(\rho_k)\mathbb{1}\{\rho_k \in [t, t+\mathrm{d}t], \boldsymbol{A}(\rho_k) \in \mathfrak{M}(\ell)\} \mid \boldsymbol{A}(0)) \\
&\quad = \mathbb{E}_{\boldsymbol{A}(0)}\left(\begin{array}{c}\mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\rho_k)\mathbb{1}\{\rho_k \in [t, t+\mathrm{d}t], \boldsymbol{A}(\rho_k) \in \mathfrak{M}(\ell)\} \mid \boldsymbol{A}(t)) \\ \times\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \text{ for all } s \leq t\}\end{array}\right) \\
&\quad = \mathbb{E}_{\boldsymbol{A}(0)}\left(\begin{array}{c}\mathbb{E}_{\boldsymbol{A}(t)}(\boldsymbol{B}(\rho_k)\mathbb{1}\{\rho_k \in [0, \mathrm{d}t], \boldsymbol{A}(\rho_k) \in \mathfrak{M}(\ell)\}) \\ \times\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \text{ for all } s \leq t\}\end{array}\right) \\
&\quad = \mathbb{E}_{\boldsymbol{A}(0)}([\boldsymbol{B}(t)\boldsymbol{\mathcal{D}}_{k\ell}\mathrm{d}t] \times \mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \text{ for all } s \leq t\}) \\
&\quad = \boldsymbol{B}(0)e^{\boldsymbol{\mathcal{C}}_k t}\boldsymbol{\mathcal{D}}_{k\ell}\mathrm{d}t.
\end{aligned}$$

3. Theorem 7.5, Remark 23.4 and Remark 23.5 imply that $\lim_{t\to\infty} e^{\boldsymbol{\mathcal{C}}_k t} = \mathbf{0}$, so that by Theorem 2.8 the eigenvalue of maximum real part of $\boldsymbol{\mathcal{C}}_k$ must have strictly negative real part. Then

$$\begin{aligned}
\mathbb{E}(\boldsymbol{B}(\rho_k)\mathbb{1}\{\boldsymbol{A}(\rho_k) \in \mathfrak{M}(\ell)\} \mid \boldsymbol{A}(0)) &= \int_0^\infty \boldsymbol{B}(0)e^{\boldsymbol{\mathcal{C}}_k t}\boldsymbol{\mathcal{D}}_{k\ell}\mathrm{d}t \\
&= \boldsymbol{B}(0)\left(\int_0^\infty e^{\boldsymbol{\mathcal{C}}_k t}\mathrm{d}t\right)\boldsymbol{\mathcal{D}}_{k\ell} \\
&= \boldsymbol{B}(0)(-\boldsymbol{\mathcal{C}}_k{}^{-1})\boldsymbol{\mathcal{D}}_{k\ell}.
\end{aligned}$$

□

7.4 Case $\mathcal{Z} = \emptyset$

In the following we study the first return and downcrossing probabilities of the process $\{(V_t, \boldsymbol{A}(t))\}_{t\geq 0}$ in the case in which $\{V_t\}_{t\geq 0}$ does not have piecewise constant intervals, that is, when $\mathcal{Z} = \emptyset$.

7.4.1 First return probabilities

Let $\boldsymbol{A}(0) \in \mathfrak{U}$ and define

$$\tau = \inf\{t > 0 : V_t \leq 0\}, \quad \Omega_\tau = \{\omega \in \Omega : \tau(\omega) < \infty\}.$$

We are interested in computing

$$\mathbb{E}_{\boldsymbol{\alpha}}(\boldsymbol{B}(\tau)\mathbb{1}(\tau < \infty));$$

for doing so, we borrow ideas from the FP3 algorithm of Bean et al. (2005).

Let Ω_1 be defined as

$$\Omega_1 = \left\{\omega = (r, a) \in \Omega_\tau : \begin{array}{c} \text{in the time interval } [0, \tau], \\ a \text{ had at most one jump from } \mathfrak{U} \text{ to } \mathfrak{D} \end{array}\right\},$$

and for $n \geq 2$, define recursively Ω_n as

$$\Omega_n = \left\{\begin{array}{l} \omega = (r, a) \in \Omega_\tau \setminus \Omega_1 : \\ \\ \\ \end{array} \begin{array}{c} \text{in the interval } [0, \tau], \text{ for } r \text{ there exist exactly} \\ \text{two succesive subexcursions above } p \\ \text{corresponding to the level process of, say,} \\ \omega^1 = \{\omega_t^1\}_{t=0}^{\tau_1} \in \Omega_{n-1}, \omega^2 = \{\omega_t^2\}_{t=0}^{\tau_2} \in \Omega_{n-1} \\ \text{shifted at } p \end{array}\right\} \cup \Omega_1, \tag{7.4.1}$$

where

$$\tau_i = \inf\{t > 0 : r_t^i = 0\} \quad (i = 1, 2), \qquad \text{and}$$

$$p = p(\omega) = \inf\left\{s \geq 0 : \begin{array}{c} \text{a transition of } a \text{ from } \mathfrak{D} \text{ to } \mathfrak{U} \text{ occured} \\ \text{at level } s \text{ in the interval } [0, \tau] \end{array}\right\}.$$

Moreover, define $\kappa_1 = p$, $\kappa_2 = \kappa_1 + \tau_1$, $\kappa_3 = \kappa_2 + \tau_2$. Figure 7.3 illustrates the previous definitions. A pathwise inspection reveals that for any $n \geq 1$,

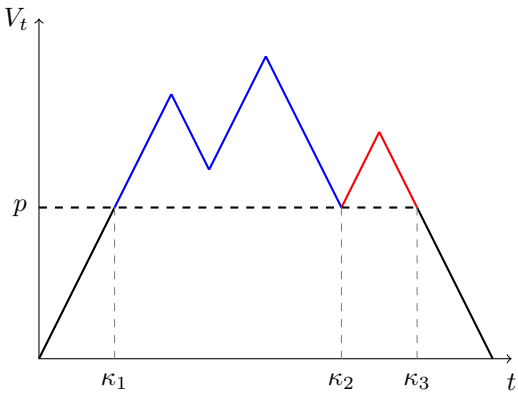


Figure 7.3: An example of a level process corresponding to $\omega \in \Omega_3$. The subexcursions above p corresponding to the level process of $\omega^1, \omega^2 \in \Omega_2$ are shown in blue and red, respectively.

$\Omega_n \subseteq \Omega_{n+1}$, and since there are no explosions of jumps in the orbit component, then

$$\Omega_\tau = \cup_{n=1}^{\infty} \Omega_n.$$

For $n \geq 1$, let

Theorem 7.7 *For all $n \geq 1$,*

$$\mathbb{E}(\boldsymbol{B}(\tau)\mathbb{1}_{\Omega_n} \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)\boldsymbol{\Psi}_n \tag{7.4.2}$$

for unique matrices $\{\Psi_n\}_{n\geq 1}$ with

$$\boldsymbol{\Psi}_0 := \mathbf{0}, \text{ and}$$

$$\boldsymbol{\Psi}_n = \int_0^\infty e^{\boldsymbol{\mathcal{C}}_+ y}(\boldsymbol{\mathcal{D}}_{+-} + \boldsymbol{\Psi}_{n-1}\boldsymbol{\mathcal{D}}_{-+}\boldsymbol{\Psi}_{n-1})e^{\boldsymbol{\mathcal{C}}_- y}\mathrm{d}y. \tag{7.4.3}$$

Proof. Proof by induction.

- **Case $n = 1$.** Let ρ_+ be as in Lemma 7.6, so that ρ_+ corresponds to the

epoch at which the first transition from $\mathfrak{U}$ to $\mathfrak{D}$ occurs. Then,

$$\begin{aligned}
\mathbb{E}(\boldsymbol{B}(\tau)\mathbb{1}_{\Omega_1} \mid \boldsymbol{A}(0)) &= \int_0^\infty \mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\tau)\mathbb{1}\{\Omega_1, \rho_+ \in [y, y+\mathrm{d}y]\}) \\
&= \int_0^\infty \mathbb{E}_{\boldsymbol{A}(0)} \left(\begin{array}{c} \mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\tau)\mathbb{1}\{\Omega_1\} \mid \rho_+, \boldsymbol{A}(\rho_+)) \\ \times \mathbb{1}\{\rho_+ \in [y, y+\mathrm{d}y]\} \end{array} \right) \\
&= \int_0^\infty \mathbb{E}_{\boldsymbol{A}(0)} \left(\begin{array}{c} \mathbb{E}_{\boldsymbol{A}(\rho_+)}(\boldsymbol{B}(y)\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{D} \text{ for all } s \le y\}) \\ \times \mathbb{1}\{\rho_+ \in [y, y+\mathrm{d}y]\} \end{array} \right) \\
&= \int_0^\infty \mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\rho_+)e^{\boldsymbol{\mathcal{C}}-y} \times \mathbb{1}\{\rho_+ \in [y, y+\mathrm{d}y]\}) \\
&= \int_0^\infty \boldsymbol{B}(0)e^{\boldsymbol{\mathcal{C}}+y}\boldsymbol{\mathcal{D}}_{+-}e^{\boldsymbol{\mathcal{C}}-y}\mathrm{d}y,
\end{aligned}$$

where the strong Markov property was used in the third equality, Theorem 7.5 was used in the second–to–last equality and (7.3.6) in the last one. Minimality of $\mathfrak{U}$ implies that

$$\boldsymbol{\Psi}_1 = \int_0^\infty e^{\boldsymbol{\mathcal{C}}+y}\boldsymbol{\mathcal{D}}_{+-}e^{\boldsymbol{\mathcal{C}}-y}\mathrm{d}y$$

is the only solution to (7.4.2) for $n = 1$.

- **Inductive part.** Suppose that (7.4.2) is true for some $n - 1 \ge 1$. First, for $V_0 = 0$, define

$$\xi = \inf\{s \ge \tau : \boldsymbol{A}(s) \in \mathfrak{U}\} \text{ and } \mu = \xi - \tau. \tag{7.4.4}$$

That is, μ corresponds to the time it takes $\boldsymbol{A}(\cdot)$ to jump to $\mathfrak{U}$ after the first downcrossing of 0 happened. Then, for any $\boldsymbol{\beta} \in \mathfrak{U}$,

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\beta}}(\boldsymbol{B}(\xi)\mathbb{1}\{\Omega_{n-1}, \mu \in [0, \mathrm{d}y]\}) &(7.4.5)\\
&\quad = \mathbb{E}_{\boldsymbol{\beta}}(\mathbb{E}_{\boldsymbol{\beta}}(\boldsymbol{B}(\xi)\mathbb{1}\{\Omega_{n-1}, \mu \in [0, \mathrm{d}y]\} \mid \{\boldsymbol{A}(s)\}_{s \le \tau})) \\
&\quad = \mathbb{E}_{\boldsymbol{\beta}}(\mathbb{1}\{\Omega_{n-1}\}\mathbb{E}_{\boldsymbol{A}(\tau)}(\boldsymbol{B}(\rho_+)\mathbb{1}\{\rho_+ \in [0, \mathrm{d}y]\})) \\
&\quad = \mathbb{E}_{\boldsymbol{\beta}}(\mathbb{1}\{\Omega_{n-1}\}[\boldsymbol{A}(\tau)\boldsymbol{\mathcal{D}}_{-+}\mathrm{d}y]) \\
&\quad = \mathbb{E}_{\boldsymbol{\beta}}(\boldsymbol{B}(0))\boldsymbol{\Psi}_{n-1}\boldsymbol{\mathcal{D}}_{-+}\mathrm{d}y, &(7.4.6)
\end{aligned}$$

where the strong Markov property, (7.3.6) and the induction hypothesis were used, in that order. To ease notation, define the events

$$\begin{aligned}
E_1^p &= \{\boldsymbol{A}(s) \in \mathfrak{U} \text{ for all } s \le p\} \\
E_2^p &= \{\text{Subexcursion } \omega^1 \text{ shifted at } p \text{ is in } \Omega_{n-1}\} \\
E_3^p &= \{\text{Subexcursion } \omega^2 \text{ shifted at } p \text{ is in } \Omega_{n-1}\}.
\end{aligned}$$

Then,

$$\begin{aligned}
&\mathbb{E}\left(\boldsymbol{B}(\tau)\mathbb{1}\{\Omega_n \setminus \Omega_1\} \mid \boldsymbol{A}(0)\right) \\
&= \int_0^\infty \mathbb{E}_{\boldsymbol{A}(0)}\left(\boldsymbol{B}(\tau)\mathbb{1}\{\Omega_n \setminus \Omega_1, p \in [y, y+\mathrm{d}y]\}\right) \\
&= \int_0^\infty \mathbb{E}_{\boldsymbol{A}(0)}\left(\begin{array}{c}\mathbb{E}_{\boldsymbol{A}(0)}\left(\boldsymbol{B}(\tau)\mathbb{1}\{\Omega_n \setminus \Omega_1\} \mid \{\boldsymbol{A}(s)\}_{s\le\kappa_3}, p\right) \\ \times\mathbb{1}\{p \in [y, y+\mathrm{d}y]\}\end{array}\right) \\
&= \int_0^\infty \mathbb{E}_{\boldsymbol{A}(0)}\left(\begin{array}{c}\mathbb{E}_{\boldsymbol{A}(\kappa_3)}\left(\boldsymbol{B}(y)\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{D} \text{ for all } s \le y\}\right) \\ \times\mathbb{1}\{E_1^p \cap E_2^p \cap E_3^p, p \in [y, y+\mathrm{d}y]\}\end{array}\right) \\
&= \int_0^\infty \mathbb{E}_{\boldsymbol{A}(0)}\left(\left[\boldsymbol{B}(\kappa_3)e^{\boldsymbol{\mathcal{C}}-y}\right]\mathbb{1}\{E_1^p \cap E_2^p \cap E_3^p, p \in [y, y+\mathrm{d}y]\}\right), \quad (7.4.7)
\end{aligned}$$

where the strong Markov property and Theorem 7.5 were used in the last two equalities. Furthermore,

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{A}(0)}\left(\boldsymbol{B}(\kappa_3)\mathbb{1}\{E_1^p \cap E_2^p \cap E_3^p, p \in [y, y+\mathrm{d}y]\}\right) \\
&\quad= \mathbb{E}_{\boldsymbol{A}(0)}\left(\begin{array}{c}\mathbb{E}_{\boldsymbol{A}(0)}\left(\boldsymbol{B}(\kappa_3)\mathbb{1}\{E_3^p\} \mid \{\boldsymbol{A}(s)\}_{s\le\kappa_2}, p\right) \\ \times\mathbb{1}\{E_1^p \cap E_2^p, p \in [y, y+\mathrm{d}y]\}\end{array}\right) \\
&\quad= \mathbb{E}_{\boldsymbol{A}(0)}\left(\begin{array}{c}\mathbb{E}_{\boldsymbol{A}(\kappa_2)}\left(\boldsymbol{B}(\tau)\mathbb{1}\{\Omega_{n-1}\}\right) \\ \times\mathbb{1}\{E_1^p \cap E_2^p, p \in [y, y+\mathrm{d}y]\}\end{array}\right) \\
&\quad= \mathbb{E}_{\boldsymbol{A}(0)}\left(\left[\boldsymbol{B}(\kappa_2)\boldsymbol{\Psi}_{n-1}\right]\mathbb{1}\{E_1^p \cap E_2^p, p \in [y, y+\mathrm{d}y]\}\right), \quad (7.4.8)
\end{aligned}$$

where induction hypothesis is used in the last equality, and

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{A}(0)}\left(\boldsymbol{B}(\kappa_2)\mathbb{1}\{E_1^p \cap E_2^p, p \in [y, y+\mathrm{d}y]\}\right) \\
&\quad= \mathbb{E}_{\boldsymbol{B}(0)}\left(\begin{array}{c}\mathbb{E}_{\boldsymbol{B}(0)}\left(\boldsymbol{B}(\kappa_2)\mathbb{1}\{E_2^p, p \in [y, y+\mathrm{d}y]\} \mid \{\boldsymbol{A}(s)\}_{s\le y}\right) \\ \times\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{U} \text{ for all } s \le y\}\end{array}\right) \\
&\quad= \mathbb{E}_{\boldsymbol{B}(0)}\left(\begin{array}{c}\mathbb{E}_{\boldsymbol{A}(y)}(\boldsymbol{B}(\xi)\mathbb{1}\{\Omega_{n-1}, \mu \in [0, \mathrm{d}y]\}) \\ \times\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{U} \text{ for all } s \le y\}\end{array}\right) \\
&\quad= \mathbb{E}_{\boldsymbol{B}(0)}\left(\left[\boldsymbol{B}(y)\boldsymbol{\Psi}_{n-1}\boldsymbol{\mathcal{D}}_{-+}\mathrm{d}y\right] \times \mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{U} \text{ for all } s \le y\}\right) \\
&\quad= \boldsymbol{B}(0)e^{\boldsymbol{\mathcal{C}}+y}\boldsymbol{\Psi}_{n-1}\boldsymbol{\mathcal{D}}_{-+}\mathrm{d}y, \quad (7.4.9)
\end{aligned}$$

where the Markov property, (7.4.6) and Theorem 7.5 were used in the last three equalities. Combining (7.4.7), (7.4.8) and (7.4.9) we get that

$$\begin{aligned}
&\mathbb{E}\left(\boldsymbol{B}(\tau)\mathbb{1}\{\Omega_n \setminus \Omega_1\} \mid \boldsymbol{A}(0)\right) \\
&\quad= \int_0^\infty \boldsymbol{B}(0)e^{\boldsymbol{\mathcal{C}}+y}\boldsymbol{\Psi}_{n-1}\boldsymbol{\mathcal{D}}_{-+}\boldsymbol{\Psi}_{n-1}e^{\boldsymbol{\mathcal{C}}-y}\mathrm{d}y.
\end{aligned}$$

Uniqueness is again implied by the minimality of $\mathfrak{U}$ and thus (7.4.3) follows.

□

Corollary 7.8 *Compute $\boldsymbol{\Psi}_n$, $n \geq 1$, recursively by setting $\boldsymbol{\Psi}_0 = \mathbf{0}$ and solving*

$$\boldsymbol{\mathcal{C}}_+ \boldsymbol{\Psi}_{n+1} + \boldsymbol{\Psi}_{n+1} \boldsymbol{\mathcal{C}}_- = -\boldsymbol{\mathcal{D}}_{+-} - \boldsymbol{\Psi}_n \boldsymbol{\mathcal{D}}_{-+} \boldsymbol{\Psi}_n. \tag{7.4.10}$$

Then $\mathbb{E}(\boldsymbol{B}(\tau)\mathbb{1}_{\Omega_\tau} \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)\boldsymbol{\Psi}$ where $\boldsymbol{\Psi} = \lim_{n\to\infty} \boldsymbol{\Psi}_n$. Furthermore,

$$\mathbb{P}(\tau < \infty \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)\boldsymbol{\Psi}\boldsymbol{e}. \tag{7.4.11}$$

Proof. In the proof of (7.3.7) we checked that the eigenvalue of maximal real part of $\boldsymbol{\mathcal{C}}_-$ (ditto $\boldsymbol{\mathcal{C}}_+$) has strictly negative real part. Thus, premultyplying (7.4.3) by $\boldsymbol{\mathcal{C}}_+$ and integrating by parts we get that

$$\begin{aligned}
\boldsymbol{\mathcal{C}}_+ \boldsymbol{\Psi}_{n+1} &= \int_0^\infty \boldsymbol{\mathcal{C}}_+ e^{\boldsymbol{\mathcal{C}}_+ y} (\boldsymbol{\mathcal{D}}_{+-} + \boldsymbol{\Psi}_n \boldsymbol{\mathcal{D}}_{-+} \boldsymbol{\Psi}_n) e^{\boldsymbol{\mathcal{C}}_- y} \mathrm{d}y \\
&= [e^{\boldsymbol{\mathcal{C}}_+ y} (\boldsymbol{\mathcal{D}}_{+-} + \boldsymbol{\Psi}_n \boldsymbol{\mathcal{D}}_{-+} \boldsymbol{\Psi}_n) e^{\boldsymbol{\mathcal{C}}_- y}]_0^\infty \\
&\quad - \int_0^\infty e^{\boldsymbol{\mathcal{C}}_+ y} (\boldsymbol{\mathcal{D}}_{+-} + \boldsymbol{\Psi}_n \boldsymbol{\mathcal{D}}_{-+} \boldsymbol{\Psi}_n) e^{\boldsymbol{\mathcal{C}}_- y} \boldsymbol{\mathcal{C}}_- \mathrm{d}y \\
&= [\mathbf{0} - (\boldsymbol{\mathcal{D}}_{+-} + \boldsymbol{\Psi}_n \boldsymbol{\mathcal{D}}_{-+} \boldsymbol{\Psi}_n)] - \boldsymbol{\Psi}_{n+1} \boldsymbol{\mathcal{C}}_-,
\end{aligned}$$

so that (7.4.10) follows. Existence of $\lim_{n\to\infty} \boldsymbol{\Psi}_n$ is implied by the boundedness of $\mathfrak{U} \cup \mathfrak{D}$ and the minimality property $\mathfrak{U}$. Equation (7.4.11) follows by noticing that since $\mathfrak{D}$ is an affine subspace, then $\boldsymbol{B}(t)\boldsymbol{e} = 1$ for all $t \geq 0$. □

Remark 24 *By taking limits in (7.4.10), we see that $\boldsymbol{\Psi}$ is a solution to the Riccati equation*

$$\boldsymbol{\mathcal{C}}_+ \boldsymbol{\Psi} + \boldsymbol{\Psi} \boldsymbol{\mathcal{C}}_- + \boldsymbol{\mathcal{D}}_{+-} + \boldsymbol{\Psi} \boldsymbol{\mathcal{D}}_{-+} \boldsymbol{\Psi} = \mathbf{0}. \tag{7.4.12}$$

In the case the parameters of the FRAP correspond to the ones of a fluid flow processes, there exist more algorithms to solve (7.4.12) which converge faster to $\boldsymbol{\Psi}$ than the one described in Corollary 7.8 (see Bean et al. (2005)). However, in the case of a general FRAP, the algorithm of Corollary 7.8 is the only one guaranteed to work (for now).

7.4.2 Downward record process

For $x \geq 0$, let us introduce the stopping time

$$\tau(x) := \inf\{t > 0 : V_t < -x\}.$$

For a fixed path $\omega \in \Omega_\tau$, let $x \geq 0$ be such that $\tau(x) < \infty$. On $[0, x]$, $\tau(\cdot)$ is a càdlàg nondecreasing function which is increasing with slope $+1$ in between its discontinuities. Define the **downward record process** $\{(\ell_x, \boldsymbol{O}(x))\}_{x\geq 0}$ by

$$(\ell_x, \boldsymbol{O}(x)) = \begin{cases} (V_{\tau(x)}, \boldsymbol{A}(\tau(x))) & \text{if} \quad \tau(x) < \infty, \\ (\infty, \boldsymbol{\Delta}) & \text{if} \quad \tau(x) = \infty, \end{cases}$$

where $\boldsymbol{\Delta}$ denotes a cemetery state. Furthermore, define $\sigma_0 = 0$ and for $n \geq 1$ let

$$\sigma_n = \inf\{x > 0 : \boldsymbol{A}(\tau(\sigma_{n-1}) + x) \in \mathfrak{U}\} + \sigma_{n-1}.$$

See Figure 7.4 for a visual explanation of the downward record process and of $\{\sigma_n\}_{n\geq 0}$. Let $V_x = \inf\{n \geq 1 : \sigma_n > x\} - 1$. If $V_x = n$, we say that the process **reached level** $-x$ **with** n **record downcrossings**. Notice that

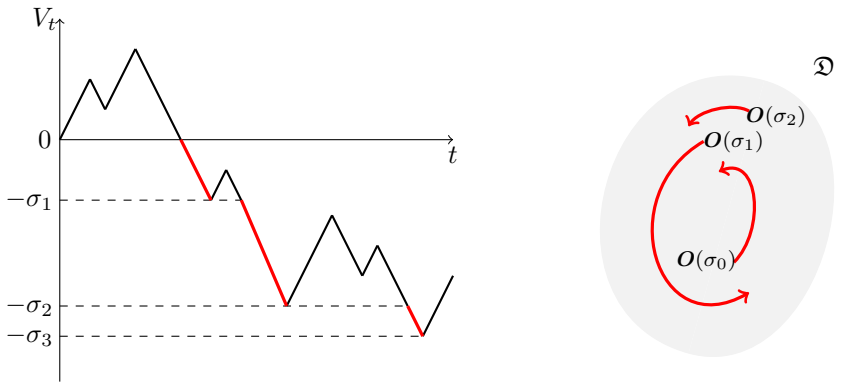


Figure 7.4: An example of the record downward process of the Proto–FRAP. On the left, record downward levels are shown in red. On the right, the concatenation $\{\boldsymbol{O}(t)\}_{t\geq 0}$ of the corresponding orbit segments is shown.

if $\tau(x) < \infty$, then $V_{\tau(x)} = -x$. We can see that $\{\boldsymbol{O}(x)\}_{x\geq 0}$ is a (possibly absorbing) concatenation of orbits with state–space $\mathfrak{D} \cup \{\boldsymbol{\Delta}\}$.

Define the process $\{(\ell_x, \boldsymbol{G}(x))\}_{x\geq 0}$ by

$$(\ell_x, \boldsymbol{G}(x)) = \begin{cases} (V_{\tau(x)}, \boldsymbol{B}(\tau(x))) & \text{if} \quad \tau(x) < \infty, \\ (\infty, \boldsymbol{\Delta}) & \text{if} \quad \tau(x) = \infty. \end{cases}$$

This is equivalent to defining

$$\boldsymbol{G}(x) = \sum_{i\in\mathcal{N}} b^{(-,i)}(\boldsymbol{O}(x)) \mathbb{1}_{\boldsymbol{O}(x)\in\mathfrak{D}_i} \quad \text{if } \tau(x) < \infty.$$

Theorem 7.9

$$\mathbb{E}(\boldsymbol{G}(x)\mathbb{1}(\tau(x) < \infty) \mid \boldsymbol{O}(0)) = \boldsymbol{G}(0)e^{(\boldsymbol{\mathcal{C}}_{-}^{*} + \boldsymbol{\mathcal{D}}_{-+}^{*}\boldsymbol{\Psi}^{*})x}. \tag{7.4.13}$$

Proof. First, we prove that for each $n \geq 0$, there exists a unique continuous matrix function $\boldsymbol{\Phi}_n(\cdot)$ such that

$$\mathbb{E}(\boldsymbol{G}(x)\mathbb{1}\{V_x = n\} \mid \boldsymbol{O}(0)) = \boldsymbol{G}(0)\boldsymbol{\Phi}_n(x). \tag{7.4.14}$$

We do this by induction.

- **Case $n = 0$.**

$$\begin{aligned}
&\mathbb{E}(\boldsymbol{G}(x)\mathbb{1}\{V_x = 0\} \mid \boldsymbol{O}(0)) \\
&\quad = \mathbb{E}_{\boldsymbol{O}(0)}(\boldsymbol{B}(x)\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{D} \text{ for all } s \leq x\}) \\
&\quad = \mathbb{E}_{\boldsymbol{O}(0)}(\boldsymbol{B}(0))e^{\boldsymbol{\mathcal{C}}_{-}x} \\
&\quad = \boldsymbol{G}(0)e^{\boldsymbol{\mathcal{C}}_{-}x}
\end{aligned}$$

  so that (7.4.14) holds with $\boldsymbol{\Phi}_0(x) = e^{\boldsymbol{\mathcal{C}}_{-}x}$. Uniqueness follows from the minimality of $\mathfrak{D}$.

- **Inductive part.** Now suppose that (7.4.14) holds for $n - 1 \geq 0$. First, notice that

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{O}(0)}(\boldsymbol{G}(\sigma_1)\mathbb{1}\{\sigma_1 \in [y, y + \mathrm{d}y]\}) \\
&\qquad = \mathbb{E}_{\boldsymbol{O}(0)}(\mathbb{E}_{\boldsymbol{O}(0)}(\boldsymbol{G}(\sigma_1)\mathbb{1}\{\sigma_1 \in [y, y + \mathrm{d}y]\} \mid \sigma_1, \boldsymbol{A}(\sigma_1))) \\
&\qquad = \mathbb{E}_{\boldsymbol{O}(0)}(\mathbb{E}_{\boldsymbol{A}(\sigma_1)}(\boldsymbol{B}(\tau)\mathbb{1}\{\tau < \infty\})\mathbb{1}\{\sigma_1 \in [y, y + \mathrm{d}y]\}) \\
&\qquad = \mathbb{E}_{\boldsymbol{O}(0)}(\boldsymbol{B}(\sigma_1)\boldsymbol{\Psi}\mathbb{1}\{\sigma_1 \in [y, y + \mathrm{d}y]\}) \\
&\qquad = \mathbb{E}_{\boldsymbol{O}(0)}(\boldsymbol{B}(\rho_{-})\mathbb{1}\{\rho_{-} \in [y, y + \mathrm{d}y]\})\boldsymbol{\Psi} \\
&\qquad = \boldsymbol{G}(0)e^{\boldsymbol{\mathcal{C}}_{-}}\boldsymbol{\mathcal{D}}_{-+}\boldsymbol{\Psi}\mathrm{d}y,
\end{aligned}$$

  where we used the strong Markov property in the second equality, Corol-

lary 7.8 in the third one and (7.3.6) in the last one. Then

$$\begin{aligned}
&\mathbb{E}(\boldsymbol{G}(x)\mathbb{1}\{V_x = n\} \mid \boldsymbol{O}(0)) \\
&\quad = \mathbb{E}_{\boldsymbol{O}(0)}(\boldsymbol{G}(x)\mathbb{1}\{V_x = n\}) \\
&\quad = \int_0^x \mathbb{E}_{\boldsymbol{O}(0)}(\boldsymbol{G}(x)\mathbb{1}\{V_x = n, \sigma_1 \in [y, y + \mathrm{d}y]\}) \\
&\quad = \int_0^x \mathbb{E}_{\boldsymbol{O}(0)}(\mathbb{E}_{\boldsymbol{O}(0)}(\boldsymbol{G}(x)\mathbb{1}\{V_x = n, \sigma_1 \in [y, y + \mathrm{d}y]\} \mid \sigma_1, \boldsymbol{O}(\sigma_1))) \\
&\quad = \int_0^x \mathbb{E}_{\boldsymbol{O}(0)}(\mathbb{E}_{\boldsymbol{O}(\sigma_1)}(\boldsymbol{G}(x - y)\mathbb{1}\{V_{x-y} = n - 1\})\mathbb{1}\{\sigma_1 \in [y, y + \mathrm{d}y]\}) \\
&\quad = \int_0^x \mathbb{E}_{\boldsymbol{O}(0)}(\boldsymbol{G}(\sigma_1)\boldsymbol{\Phi}_{n-1}(x - y)\mathbb{1}\{\sigma_1 \in [y, y + \mathrm{d}y]\}) \\
&\quad = \int_0^x \boldsymbol{G}(0)e^{\boldsymbol{\mathcal{C}}_- y}\boldsymbol{\mathcal{D}}_{-+}\boldsymbol{\Psi}\boldsymbol{\Phi}_{n-1}(x - y)\mathrm{d}y,
\end{aligned}$$

where the strong Markov property is used in the fourth equality, induction hypothesis in the second–to–last equality and (7.3.6) in the last one. Thus, (7.4.14) recursively holds with

$$\boldsymbol{\Phi}_n(x) = \int_0^x e^{\boldsymbol{\mathcal{C}}_- y}\boldsymbol{\mathcal{D}}_{-+}\boldsymbol{\Psi}\boldsymbol{\Phi}_{n-1}(x - y)\mathrm{d}y, \tag{7.4.15}$$

which happens to be unique by the minimality assumption of $\mathfrak{D}$.

Summing (7.4.15) over $n \geq 0$ and using the boundedness assumption of $\mathfrak{D}$, we get that

$$\mathbb{E}(\boldsymbol{G}(x)\mathbb{1}(\tau(x) < \infty) \mid \boldsymbol{O}(0) = \boldsymbol{\beta}) = \boldsymbol{G}(0)\boldsymbol{\Phi}(x)$$

for a unique and $\|\cdot\|_{\max}$–bounded $\boldsymbol{\Phi}(\cdot)$ satisfying

$$\boldsymbol{\Phi}(x) = e^{\boldsymbol{\mathcal{C}}_- y} + \int_0^x e^{\boldsymbol{\mathcal{C}}_- y}\boldsymbol{\mathcal{D}}_{-+}\boldsymbol{\Psi}\boldsymbol{\Phi}(x - y)\mathrm{d}y.$$

Theorem 7.4 implies that $\boldsymbol{\Phi}(x) = e^{(\boldsymbol{\mathcal{C}}_- + \boldsymbol{\mathcal{D}}_{-+}\boldsymbol{\Psi})x}$ and the proof is finished. □

Corollary 7.10 *Let* $\boldsymbol{A}(0) \in \mathfrak{U}$. *Then*

$$\mathbb{E}(\boldsymbol{G}(x)\mathbb{1}(\tau(x) < \infty) \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)\boldsymbol{\Psi}e^{(\boldsymbol{\mathcal{C}}_- + \boldsymbol{\mathcal{D}}_{-+}\boldsymbol{\Psi})x}. \tag{7.4.16}$$

Furthermore,

$$\mathbb{P}(\tau(x) < \infty \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)\boldsymbol{\Psi}\boldsymbol{e}. \tag{7.4.17}$$

PROOF. The first assertion follows from Corollary 7.8, Theorem 7.9 and from the fact that on $\{\tau < \infty\}$, $\boldsymbol{G}(0) = \boldsymbol{B}(\tau)$. Equation (7.4.17) follows by noticing that the affine nature of the state–space of $\{\boldsymbol{A}(t)\}_{t\geq 0}$ implies that on the event $\{\tau(x) < \infty\}$, $\boldsymbol{G}(x)\boldsymbol{e} = 1$. □

7.5 Case $\mathcal{Z} \neq \emptyset$

In this section we generalize the results of Section 7.4 to the case in which $\mathcal{Z} = \emptyset$. Let

$$W_t = \int_0^t \mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{U} \cup \mathfrak{D}\}\mathrm{d}s;$$

we call W_t the **total fluid up to time** t. Furthermore, for all $x \geq 0$ define

$$\zeta_x = \inf\{t \geq 0 : W_t > x\},$$

so that $\{\zeta_x\}_{x\geq 0}$ is the càdlàg process which records the amount of time it takes the total fluid to be equal to x.

Theorem 7.11 *Let $\boldsymbol{A}(0) \in \mathfrak{M}(k)$, $k \in \{+,-\}$. Then*

$$\mathbb{E}(\boldsymbol{B}(\zeta_x)\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \cup \mathfrak{Z} \text{ for all } s \leq \zeta_x\} \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)e^{\boldsymbol{\mathcal{C}}_k^* x}, \qquad (7.5.1)$$

where

$$\boldsymbol{\mathcal{C}}_k^* = \boldsymbol{\mathcal{C}}_k + \boldsymbol{\mathcal{D}}_{k0}(-\boldsymbol{\mathcal{C}}_0^{-1})\boldsymbol{\mathcal{D}}_{0k}$$

PROOF.

Let $\gamma_0 = 0$ and for $n \geq 1$

$$\gamma_n = \inf\{x > \gamma_{n-1} : \boldsymbol{A}(\zeta_{x^-}) \in \mathfrak{Z}\},$$

so that γ_n is the total fluid at which $\{\boldsymbol{A}(t)\}_{t\geq 0}$ enters $\mathfrak{Z}$ for the n-th time. Let

$$K_x := \inf\{n \geq 0 : \gamma_n < x, \boldsymbol{A}(s) \in \mathfrak{M}(k) \cup \mathfrak{Z} \text{ for all } s \leq \zeta_x\},$$

so that $\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \cup \mathfrak{Z} \text{ for all } s \leq \zeta_x\}$ if and only if $\{K_x < \infty\}$, and on this event, K_x corresponds the number of jumps from $\mathfrak{M}(k)$ to $\mathfrak{Z}$ before time ζ_x.

First, we show that for all $n \geq 0$

$$\mathbb{E}(\boldsymbol{B}(\zeta_x)\mathbb{1}\{K_x = n\} \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)\boldsymbol{\Gamma}_n(x) \qquad (7.5.2)$$

for appropiate unique continuous matrices $\{\boldsymbol{\Gamma}_n(x)\}_{x\geq 0}$ $(n \geq 0)$: this is done by induction.

1. **Case** $n = 0$**.**

$$\begin{aligned}&\mathbb{E}(\boldsymbol{B}(\zeta_x)\mathbb{1}\{K_x = 0\} \mid \boldsymbol{A}(0))\\&\quad = \mathbb{E}(\boldsymbol{B}(\zeta_x)\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \text{ for all } s \le \zeta_x\} \mid \boldsymbol{A}(0))\\&\quad = \boldsymbol{B}(0)e^{\boldsymbol{\mathcal{C}}_k x},\end{aligned}$$

where the last equality followed from Theorem 7.5. This means that $\boldsymbol{\Gamma}_0(x) = e^{\boldsymbol{\mathcal{C}}_k x}$, which is continuous in x, and it is unique because of the minimality property of $\mathfrak{M}(k)$.

2. **Inductive part.** Suppose that (7.5.2) holds for some $n \ge 0$. Then,

$$\begin{aligned}&\mathbb{E}(\boldsymbol{B}(\zeta_x)\mathbb{1}\{K_x = n+1\} \mid \boldsymbol{A}(0))\\&\quad = \int_0^x \mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\zeta_x)\mathbb{1}\{K_x = n+1, \gamma_1 \in [y, \mathrm{d}y]\})\\&\quad = \int_0^x \mathbb{E}_{\boldsymbol{A}(0)}\left(\begin{array}{c}\mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\zeta_x)\mathbb{1}\{K_x = n+1\} \mid \{\boldsymbol{A}(s)\}_{s \le \zeta_{\gamma_1}}, \gamma_1)\\ \times \mathbb{1}\{\gamma_1 \in [y, \mathrm{d}y]\}\end{array}\right)\\&\quad = \int_0^x \mathbb{E}_{\boldsymbol{A}(0)}(\mathbb{E}_{\boldsymbol{A}(\zeta_{\gamma_1})}(\boldsymbol{B}(\zeta_{x-y})\mathbb{1}\{K_{x-y} = n\}) \times \mathbb{1}\{\gamma_1 \in [y, \mathrm{d}y]\})\\&\quad = \int_0^x \mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\zeta_{\gamma_1})\boldsymbol{\Gamma}_n(x-y)\mathbb{1}\{\gamma_1 \in [y, \mathrm{d}y]\}),\end{aligned}$$

where the strong Markov property and the inductive hypothesis were used in the last steps. Now,

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\zeta_{\gamma_1})\mathbb{1}\{\gamma_1 \in [y, \mathrm{d}y]\})\\&\quad = \mathbb{E}_{\boldsymbol{A}(0)}(\mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\zeta_{\gamma_1}) \mid \boldsymbol{A}(\zeta_{\gamma_1^-}), \gamma_1) \times \mathbb{1}\{\gamma_1 \in [y, \mathrm{d}y]\})\\&\quad = \mathbb{E}_{\boldsymbol{A}(0)}(\mathbb{E}_{\boldsymbol{A}(\zeta_{\gamma_1^-})}(\boldsymbol{B}(\rho_+)\mathbb{1}\{\boldsymbol{B}(\rho_+) \in \mathfrak{M}(k)\}) \times \mathbb{1}\{\gamma_1 \in [y, \mathrm{d}y]\}))\\&\quad = \mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\zeta_{\gamma_1^-})(-\boldsymbol{\mathcal{C}}_0{}^{-1})\boldsymbol{\mathcal{D}}_{0k}\mathbb{1}\{\gamma_1 \in [y, \mathrm{d}y]\})),\end{aligned}$$

where the strong Markov property and (7.3.7) were used in the last two equalities. Furthermore, (7.3.6) implies that

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\zeta_{\gamma_1^-})\mathbb{1}\{\gamma_1 \in [y, \mathrm{d}y]\}))\\&\qquad = \mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\rho_k)\mathbb{1}\{\boldsymbol{A}(\rho_k) \in \mathfrak{Z}, \rho_k \in [y, \mathrm{d}y]\}))\\&\qquad = \boldsymbol{B}(0)e^{\boldsymbol{\mathcal{C}}_k t}\boldsymbol{\mathcal{D}}_{k0}\mathrm{d}y.\end{aligned}$$

Thus, (7.5.2) holds for $n+1$ with

$$\boldsymbol{\Gamma}_{n+1}(x) = \int_0^x e^{\boldsymbol{\mathcal{C}}_k s}\boldsymbol{\mathcal{D}}_{k\ell}(-\boldsymbol{\mathcal{C}}_0{}^{-1})\boldsymbol{\mathcal{D}}_{0k}\boldsymbol{\Gamma}_n(x-s)\mathrm{d}s, \tag{7.5.3}$$

which is clearly continuous in x and unique because of the minimality of $\mathfrak{M}(k)$. Summing (7.5.3) over $n \geq 0$ and using the boundedness assumption of $\mathfrak{M}(k)$, we get

$$\mathbb{E}(\boldsymbol{B}(\zeta_x)\mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{M}(k) \cup \mathfrak{Z} \text{ for all } s \leq \zeta_x\} \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)\boldsymbol{\Gamma}(x),$$

where $\boldsymbol{\Gamma}(\cdot)$ is $\|\cdot\|_{\max}$–bounded and it is such that

$$\boldsymbol{\Gamma}(x) = e^{\boldsymbol{\mathcal{C}}_k x} + \int_0^x e^{\boldsymbol{\mathcal{C}}_k s}\boldsymbol{\mathcal{D}}_{k\ell}(-\boldsymbol{\mathcal{C}}_0{}^{-1})\boldsymbol{\mathcal{D}}_{0k}\boldsymbol{\Gamma}(x-s)\mathrm{d}s.$$

Theorem 7.4 implies that $\boldsymbol{\Gamma}(x) = e^{(\boldsymbol{\mathcal{C}}_k + \boldsymbol{\mathcal{D}}_{k0}(-\boldsymbol{\mathcal{C}}_0^{-1})\boldsymbol{\mathcal{D}}_{0k})x}$ and (7.5.1) follows.

□

Lemma 7.12 *Let $k \in \{+,-\}$, $\boldsymbol{A}(0) \in \mathfrak{M}(k)$. Then,*

1. *We have that*

$$\mathbb{P}(\boldsymbol{A}(s) \in \mathfrak{M}(k) \cup \mathfrak{Z} \text{ for all } s \leq \zeta_x \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)e^{\boldsymbol{\mathcal{C}}_k^* x}\boldsymbol{e}.$$

2. *Let $\ell = \{+,-,0\} \setminus \{0,k\}$ and define*

$$\delta_\ell = \inf\{x > 0 : \boldsymbol{A}(\zeta_x) \in \mathfrak{M}(\ell)\},$$

so that δ_ℓ is the total fluid at which $\{\boldsymbol{A}(t)\}_{t\geq 0}$ enters $\mathfrak{M}(\ell)$ for the first time. Then

$$\mathbb{E}(\boldsymbol{B}(\zeta_{\delta_\ell})\mathbb{1}\{\delta_\ell \in [x, x+\mathrm{d}x]\} \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)e^{\boldsymbol{\mathcal{C}}_k^* x}\boldsymbol{\mathcal{D}}_{k\ell}^*\mathrm{d}x. \tag{7.5.4}$$

where

$$\boldsymbol{\mathcal{D}}_{k\ell}^* := \boldsymbol{\mathcal{D}}_{k\ell} + \boldsymbol{\mathcal{D}}_{k0}(-\boldsymbol{\mathcal{C}}_0^{-1})\boldsymbol{\mathcal{D}}_{0\ell}.$$

Proof.

1. This follows directly from Theorem 7.11.
2. Notice that the first jump to $\mathfrak{M}(\ell)$ at ζ_{δ_ℓ} can happen in two ways: either it comes from $\mathfrak{Z}$ (which happens if $\boldsymbol{A}(\zeta_{\delta_k^-}) \in \mathfrak{Z}$), or it comes from $\mathfrak{M}(k)$ (which happens if $\zeta_{\delta_k^-} = \zeta_{\delta_k}$). See Figure 7.5. Thus,

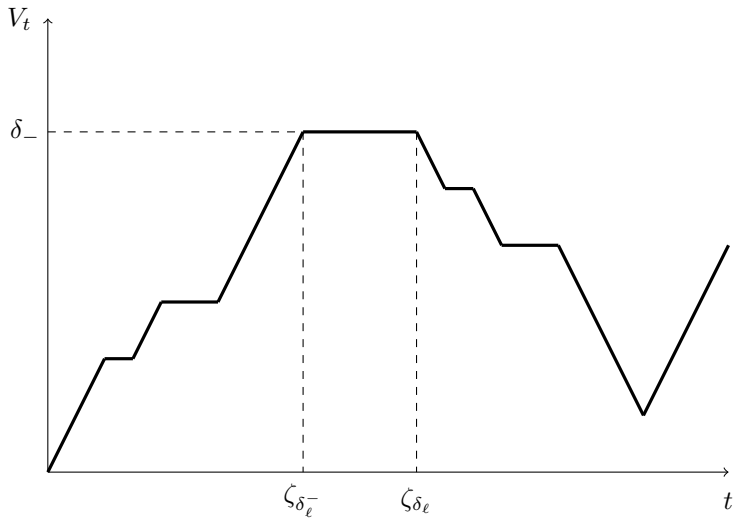


Figure 7.5: A level process in which, for $\ell = -$, $\zeta_{\delta_\ell^-} \neq \zeta_{\delta_\ell^-}$ and $\boldsymbol{A}(\zeta_{\delta_\ell^-}) \in \mathfrak{Z}$ while $\boldsymbol{A}(\zeta_{\delta_\ell}) \in \mathfrak{J}(\ell)$.

$$
\begin{aligned}
&\mathbb{E}(\boldsymbol{B}(\zeta_{\delta_\ell})\mathbb{1}\{\delta_\ell \in [y, y+\mathrm{d}y]\} \mid \boldsymbol{A}(0)) \\
&\quad = \mathbb{E}_{\boldsymbol{A}(0)}(\mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\zeta_{\delta_\ell})\mathbb{1}\{\zeta_{\delta_\ell^-} = \zeta_{\delta_\ell}, \delta_\ell \in [y, y+\mathrm{d}y]\} \mid \boldsymbol{A}(\zeta_{\delta_y}))) \\
&\qquad + \mathbb{E}_{\boldsymbol{A}(0)}(\mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\zeta_{\delta_\ell}) \times \mathbb{1}\{\boldsymbol{A}(\zeta_{\delta_\ell^-}) \in \mathfrak{Z}, \delta_\ell \in [y, y+\mathrm{d}y]\} \mid \boldsymbol{A}(\zeta_{\delta_\ell^-}), \delta_\ell)) \\
&\quad = \mathbb{E}_{\boldsymbol{A}(0)} \begin{pmatrix} \mathbb{E}_{\boldsymbol{A}(\zeta_y)}(\boldsymbol{B}(\rho_k)\mathbb{1}\{\boldsymbol{A}(\rho_k) \in \mathfrak{M}(\ell), \rho_k \in [0, \mathrm{d}y]\}) \\ \times \mathbb{1}\{\boldsymbol{A}(s) \text{ for all } s \leq \zeta_y\} \end{pmatrix} \\
&\qquad + \mathbb{E}_{\boldsymbol{A}(0)} \begin{pmatrix} \mathbb{E}_{\boldsymbol{B}(\zeta_{\delta_\ell^-})}(\boldsymbol{B}(\rho_0)\mathbb{1}\{\boldsymbol{A}(\rho_0) \in \mathfrak{M}(\ell)\}) \\ \times \mathbb{1}\{\boldsymbol{A}(\zeta_{\delta_\ell^-}) \in \mathfrak{Z}, \delta_\ell \in [y, y+\mathrm{d}y]\} \end{pmatrix} \\
&\quad = \mathbb{E}_{\boldsymbol{A}(0)}([\boldsymbol{B}(\zeta_y)\boldsymbol{\mathcal{D}}_{k\ell}\mathrm{d}y] \times \mathbb{1}\{\boldsymbol{A}(s) \text{ for all } s \leq \zeta_y\}) \\
&\qquad + \mathbb{E}_{\boldsymbol{A}(0)}([\boldsymbol{B}(\zeta_{\delta_\ell^-})(-\boldsymbol{\mathcal{C}}_0^{-1})\boldsymbol{\mathcal{D}}_{0\ell}] \times \mathbb{1}\{\boldsymbol{A}(\zeta_{\delta_\ell^-}) \in \mathfrak{Z}, \delta_\ell \in [y, y+\mathrm{d}y]\}) \\
&\quad = \boldsymbol{A}(0)e^{\boldsymbol{\mathcal{C}}_k^* y}\boldsymbol{\mathcal{D}}_{k\ell}\mathrm{d}y \\
&\qquad + \mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\zeta_{\delta_\ell^-})\mathbb{1}\{\boldsymbol{A}(\zeta_{\delta_\ell^-}) \in \mathfrak{Z}, \delta_\ell \in [y, y+\mathrm{d}y]\})(-\boldsymbol{\mathcal{C}}_0^{-1})\boldsymbol{\mathcal{D}}_{0\ell}, \quad (7.5.5)
\end{aligned}
$$

where the strong Markov property was used in the second equality, (7.3.6) and (7.3.7) in the third one, and Theorem 7.11 in the last one. With

similar arguments, we get that

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\zeta_{\delta_\ell^-})\mathbb{1}\{\boldsymbol{A}(\zeta_{\delta_\ell^-}) \in \mathfrak{Z}, \delta_\ell \in [y, y+\mathrm{d}y]\}) \\
&\quad = \mathbb{E}_{\boldsymbol{A}(0)}(\mathbb{E}_{\boldsymbol{A}(0)}(\boldsymbol{B}(\zeta_{\delta_\ell^-})\mathbb{1}\{\boldsymbol{A}(\zeta_{\delta_\ell^-}) \in \mathfrak{Z}, \delta_\ell \in [y, y+\mathrm{d}y]\} \mid \boldsymbol{A}(\zeta_y))) \\
&\quad = \mathbb{E}_{\boldsymbol{A}(0)}\left(\begin{array}{c} \mathbb{E}_{\boldsymbol{A}(\zeta_y)}(\boldsymbol{B}(\rho_k)\mathbb{1}\{\boldsymbol{A}(\rho_+) \in \mathfrak{Z}, \rho_+ \in [0, \mathrm{d}y]\}) \\ \times \mathbb{1}\{\boldsymbol{A}(s) \text{ for all } s \leq \zeta_y\} \end{array} \right) \\
&\quad = \mathbb{E}_{\boldsymbol{A}(0)}([\boldsymbol{B}(\zeta_y)\boldsymbol{\mathcal{D}}_{k0}\mathrm{d}y] \times \mathbb{1}\{\boldsymbol{A}(s) \text{ for all } s \leq \zeta_y\}) \\
&\quad = \boldsymbol{B}(0)e^{\boldsymbol{\mathcal{C}}_k^* y}\boldsymbol{\mathcal{D}}_{k\ell}\mathrm{d}y, \qquad (7.5.6)
\end{aligned}$$

so combining (7.5.5) and (7.5.6), the result follows.

□

Now, define Let

$$\Omega_1^* = \left\{ \omega = (r, a) \in \Omega_\tau : \begin{array}{c} \text{in the time interval } [0, \tau], \\ a \text{ had at most one jump from } \mathfrak{U} \cup \mathfrak{Z} \text{ to } \mathfrak{D} \end{array} \right\},$$

and for $n \geq 2$ define recursively

$$\Omega_n^* = \left\{ \omega = (r, a) \in \Omega_\tau \setminus \Omega_1^* : \begin{array}{c} \text{in the interval } [0, \tau], \text{ for } r \text{ there exist exactly} \\ \text{two succesive subexcursions above } p \\ \text{corresponding to the level process of, say,} \\ \omega^1 = \{\omega_t^1\}_{t=0}^{\tau_1} \in \Omega_{n-1}^*, \omega^2 = \{\omega_t^2\}_{t=0}^{\tau_2} \in \Omega_{n-1}^* \\ \text{shifted at } p^* \end{array} \right\} \cup \Omega_1, \qquad (7.5.7)$$

where

$$p^* = p^*(\omega) = \inf \left\{ s \geq 0 : \begin{array}{c} \text{a transition of } a \text{ from either } \mathfrak{D} \text{ to } \mathfrak{U}, \text{ or } \mathfrak{D} \text{ to} \\ \mathfrak{Z} \text{ to } \mathfrak{U} \text{ occured at level } s \text{ in the interval } [0, \tau] \end{array} \right\}.$$

Moreover, define $\kappa_1^* = p^*$, $\kappa_2^* = \inf\{s \geq \kappa_1^* + \tau_1 : \boldsymbol{A}(s) \in \mathfrak{U}\}$, $\kappa_3 = \kappa_2^* + \tau_2$. Figure 7.6 illustrates the previous definitions.

By repeating step by step the results of Section 7.4.1 and Section 7.4.2, replacing the use of Theorem 7.5 and Lemma 7.6 with Theorem 7.11 and Lemma 7.12, we can prove the following.

Theorem 7.13 *Let $\boldsymbol{A}(0) \in \mathfrak{U}$. Then Then $\mathbb{E}(\boldsymbol{B}(\tau)\mathbb{1}_{\Omega_\tau} \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)\boldsymbol{\Psi}^*$ where $\boldsymbol{\Psi}^* = \lim_{n\to\infty} \boldsymbol{\Psi}_n^*$ and $\{\boldsymbol{\Psi}_n^*\}_{n\geq 0}$ are recursively computed by setting $\boldsymbol{\Psi}_0^* = \mathbf{0}$ and solving*

$$\boldsymbol{\mathcal{C}}_+^* \boldsymbol{\Psi}_{n+1}^* + \boldsymbol{\Psi}_{n+1}^* \boldsymbol{\mathcal{C}}_-^* = -\boldsymbol{\mathcal{D}}_{+-}^* - \boldsymbol{\Psi}_n^* \boldsymbol{\mathcal{D}}_{-+}^* \boldsymbol{\Psi}_n^*. \qquad (7.5.8)$$

Furthermore,

$$\mathbb{E}(\boldsymbol{G}(x)\mathbb{1}(\tau(x) < \infty) \mid \boldsymbol{A}(0)) = \boldsymbol{B}(0)\boldsymbol{\Psi}^* e^{(\boldsymbol{\mathcal{C}}_-^* + \boldsymbol{\mathcal{D}}_{-+}^* \boldsymbol{\Psi}^*)x}. \qquad (7.5.9)$$

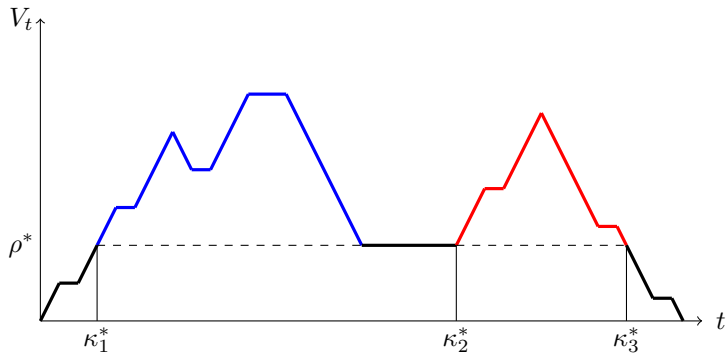


Figure 7.6: An example of a level process corresponding to $\omega \in \Omega_3^*$. The subexcursions above p^* corresponding to the level process of $\omega^1, \omega^2 \in \Omega_2^*$ are shown in blue and red, respectively.

7.6 Numerical examples.

Now, we compute the downcrossing probabilities of a FRAP that is not a fluid flow process. Its construction is based on the RAP of Example 2.6, which we describe next.

Let $\mathfrak{A}_0 := \{(1-2a, a, a) : a \in [-0.95, -0.74]\}$. For any initial point $\boldsymbol{\alpha}_0 \in \mathfrak{A}_0$, the pair $(\boldsymbol{C}, \boldsymbol{D})$ with

$$\boldsymbol{C} = \begin{pmatrix} -1 & 0 & 0 \\ -2/3 & -1 & 1 \\ 2/3 & -1 & -1 \end{pmatrix}, \qquad \text{and} \qquad \boldsymbol{D} = \begin{pmatrix} 14/5 & -9/10 & -9/10 \\ 26/15 & -8/15 & -8/15 \\ 58/15 & -19/15 & -19/15 \end{pmatrix}$$

corresponds to a RAP with jumps that always land in $\mathfrak{A}_0 \subset \mathfrak{A}$, where $\mathfrak{A}$ denotes the state–space of this RAP.

Now, let $\mathfrak{U} = \mathfrak{A}$ and let $\mathfrak{D}$ be a copy of $\mathfrak{A}$ (regarded as a different set space). For any fixed $\boldsymbol{\alpha_0} \in \mathfrak{A}_0$, consider the FRAP with state–space $\mathfrak{U} \cup \mathfrak{D}$, initial point $\boldsymbol{\alpha_0} \in \mathfrak{A}_0$, and parameters

$$\boldsymbol{C}_+ = \boldsymbol{C}_- = \boldsymbol{C}, \quad \boldsymbol{D}_{+-} = \boldsymbol{D} \quad \text{and} \quad \boldsymbol{D}_{-+} = -\boldsymbol{C}\boldsymbol{e}\boldsymbol{\alpha}_0.$$

Since $\boldsymbol{C}_+ = \boldsymbol{C}_-$, then the jump intensities and the continuous movement of the orbit in $\mathfrak{U}$ behaves exactly the same as if it were in $\mathfrak{D}$. The only difference is that a jump from happening from, say, $\boldsymbol{a} \in \mathfrak{U}$ goes to $\boldsymbol{aD}/\boldsymbol{aDe} \in \mathfrak{D}$, while jumps from $\mathfrak{D}$ always land in $\boldsymbol{\alpha}_0 \in \mathfrak{U}$. Indeed, if a jump happens, say at $\boldsymbol{b} \in \mathfrak{D}$,

it will land in

$$\frac{\boldsymbol{b}\boldsymbol{D}_{-+}}{\boldsymbol{b}\boldsymbol{D}_{-+}\boldsymbol{e}} = \frac{\boldsymbol{b}(-\boldsymbol{C}\boldsymbol{e}\boldsymbol{\alpha}_0)}{\boldsymbol{b}(-\boldsymbol{C}\boldsymbol{e}\boldsymbol{\alpha}_0)\boldsymbol{e}} = \frac{(-\boldsymbol{b}\boldsymbol{C}\boldsymbol{e})\boldsymbol{\alpha}_0}{(-\boldsymbol{b}\boldsymbol{C}\boldsymbol{e})(\boldsymbol{\alpha}_0\boldsymbol{e})} = \frac{\boldsymbol{\alpha}_0}{\boldsymbol{\alpha}_0\boldsymbol{e}} = \boldsymbol{\alpha}_0 \in \mathfrak{U},$$

which is independent from $\boldsymbol{b}$.

Now, let us choose $\boldsymbol{\alpha}_0 = (2.9, -0.95, -0.95)$, so that

$$\boldsymbol{D}_{-+} = -\boldsymbol{C}\boldsymbol{e}\boldsymbol{\alpha}_0 = \begin{pmatrix} 14.5/5 & -9.5/10 & -9.5/10 \\ 29/15 & -9.5/15 & -9.5/15 \\ 58/15 & -19/15 & -19/15 \end{pmatrix}.$$

Following the algorithm of Corollary 7.8, we get that

$$\boldsymbol{\Psi} = \begin{pmatrix} 1.8152 & -0.0979 & -0.7406 \\ 1.6990 & -0.0150 & -0.7071 \\ 2.1041 & -0.2335 & -0.8864 \end{pmatrix} \qquad \text{with} \qquad \boldsymbol{\alpha}_0\boldsymbol{\Psi}\boldsymbol{e} = 0.9694.$$

The strict inequality in $\boldsymbol{\alpha}_0\boldsymbol{\Psi}\boldsymbol{e} < 1$ implies that $\lim_{t\to\infty} V_t = +\infty$. Its downward record probabilities, $\mathbb{P}_{\boldsymbol{\alpha}_0}(\tau(x) < \infty) = \boldsymbol{\alpha}_0\boldsymbol{\Psi}e^{(\boldsymbol{C}_{-+}\boldsymbol{D}_{-+}\boldsymbol{\Psi})x}\boldsymbol{e}$ are shown in Figure 7.7.

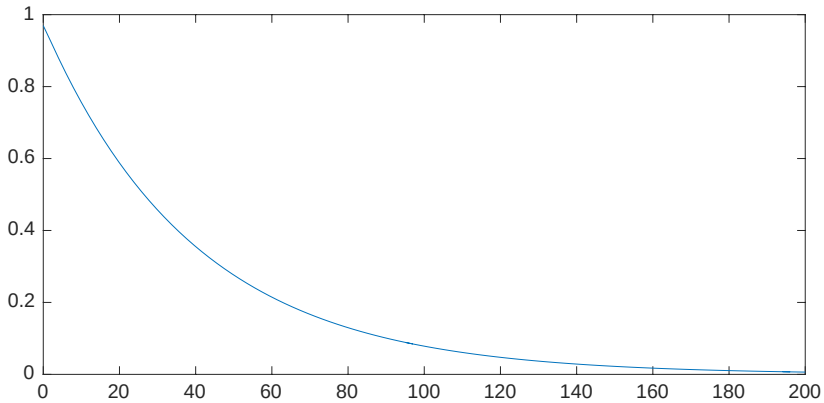


Figure 7.7: Downward record probabilities $\mathbb{P}_{\boldsymbol{\alpha}_0}(\tau(x) < \infty)$ of the example Proto–FRAP when $\boldsymbol{\alpha}_0 = (2.9, -0.95, -0.95)$.

7.7 Conclusions and remarks

In this chapter we constructed and studied first passage probabilities of the Fluid RAP (FRAP), a generalisation of the fluid flow process without Brownian

components. Below is a brief analysis of our findings and their comparison with the existing literature.

- Some attempts to study models with RAP components have been made in the literature. For instance, a study of Stochastic Petri Nets with matrix–exponential (ME) firing times was initiated in Buchholz and Telek (2010), later to be generalized in Buchholz and Telek (2013) for a class of Stochastic Automata Networks with RAP components. In Bean and Nielsen (2010), the authors construct a Quasi–Birth–Death (QBD) process with RAP components based on the orbit physical interpretation of Asmussen and Bladt (1999). Our construction of the FRAP is somewhat similar to how the QBD–RAP was introduced in Bean and Nielsen (2010), however, the methods we used to rigorously study the probabilistic properties of the path of a FRAP are completely different.
- We proved the existence of a matrix $\boldsymbol{\Psi}^*$ that characterises the first return probabilities of the FRAP. In the case of fluid flow processes, the existence of $\boldsymbol{\Psi}^*$ is trivial, since $(\boldsymbol{\Psi}^*)_{ij}$ simply corresponds to the probability that a process that starts in state i and level 0, returns to level 0 while in state j. In the case of the FRAP $(\boldsymbol{\Psi}^*)_{ij}$ has no meaning at all. In fact, the matrix $\boldsymbol{\Psi}^*$ only makes sense as a physical object **after** being pre–multiplied by an element of $\mathfrak{U}$. Moreover, the existence of $\boldsymbol{\Psi}^*$ needed to be proved while simultaneously proving that an algorithm to compute it existed.
- We also proved the existence of a matrix that characterises its downward record process, given by $\boldsymbol{\mathcal{C}}^*_- + \boldsymbol{\mathcal{D}}^*_{-+}\boldsymbol{\Psi}^*$. In the case of the fluid flow process, the matrix $\boldsymbol{\mathcal{C}}^*_- + \boldsymbol{\mathcal{D}}^*_{-+}\boldsymbol{\Psi}^*$ corresponds to the intensity matrix of the projected downward record process. In the FRAP case, the matrix $\boldsymbol{\mathcal{C}}^*_- + \boldsymbol{\mathcal{D}}^*_{-+}\boldsymbol{\Psi}^*$ has no physical or probabilistic meaning by itself.
- Even with orbit physical interpretations available, the main challenge encountered when studying models with ME or RAP components (such as the FRAP) is that one cannot make an entrywise study of the vectors and matrices involved, since they only make sense as a whole. A large portion of the arguments in the PH and MAP literature fail to do this, meaning that one needs to use different techniques to prove those statements when the models have ME or RAP components instead. It is likely that ME and RAP results will look the same as their PH and MAP counterparts, however, it is the way to get those results what makes their study interesting.

Chapter 8

Perspectives

As a conclusion, we list some research perspectives which are in the continuation of the work presented in this thesis.

- **MPH* copula.** In Chapter 4 we constructed a bivariate distribution with given phase–type marginals, which is based on the use of Baker's copula. That is, we used ideas underlying an specific copula to construct certain subclass of MPH* distributions. A further research topic is to study the class of copulas induced by MPH* distributions and how it fits within the copula theory framework.
- **Probabilistic interpretation of the Wiener–Hopf factorisation at an ME horizon.** In Chapter 6 we give a Wiener–Hopf factorisation formula for a spectrally negative Lévy process which is inspected at a ME–distributed time, using functional calculus. However, the probabilistic interpretation of such a formula is still lacking. Such an interpretation would help us to study more complex fluctuation problems, for instance, Parisian ruin with ME–distributed clocks.
- **FRAP with Brownian noise**. In Chapter 7, we constructed the FRAP as a generalisation of the fluid flow process without Brownian motion. The next step is to consider a FRAP with Brownian noise. A key element to its analysis is the Wiener–Hopf factorisation explained in the previous item, which is unavailable at the moment.

- **FRAP queue.** An interesting problem is to study the stationary distribution of a FRAP whose level is reflected at 0. Time–reversal arguments can be used in the fluid flow process case, however, those do not work for the FRAP. Alternative approaches are currently being investigated.

Appendix A

Kronecker product and Kronecker sum

Fix $n, m, r, s \geq 1$. Let $\boldsymbol{A} = \{a_{ij}\}$ be a $m \times n$–dimensional matrix and let $\boldsymbol{B} = \{b_{ij}\}$ be a $r \times s$–dimensional matrix. The **Kronecker product** of $\boldsymbol{A}$ and $\boldsymbol{B}$, denoted by $\boldsymbol{A} \otimes \boldsymbol{B}$, is defined by

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \cdots & a_{2n}\boldsymbol{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\boldsymbol{B} & a_{m2}\boldsymbol{B} & \cdots & a_{mn}\boldsymbol{B} \end{pmatrix}.$$

Thus, $\boldsymbol{A} \otimes \boldsymbol{B}$ is a $mr \times ns$–dimensional matrix. The following are some particularly useful properties of the Kronecker product.

Proposition A.1 *For matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ and $\boldsymbol{D}$, it follows that*

$$(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = \boldsymbol{AC} \otimes \boldsymbol{BD}$$

whenever the dimensions of the matrices allow it.

Proof. See Section 2.3.VI in Graham (2018).

Proposition A.2 *For any square matrix* $\boldsymbol{A}$ *and identity matrix* I *of any dimension, it follows that*

$$e^{\boldsymbol{A}\otimes\boldsymbol{I}} = e^{\boldsymbol{A}} \otimes \boldsymbol{I} \quad \textit{and} \quad e^{\boldsymbol{I}\otimes\boldsymbol{A}} = \boldsymbol{I} \otimes e^{\boldsymbol{A}}.$$

Proof. Follows by the series expansion of the matrix exponential terms.

Now, for $n, m \geq 1$ let $\boldsymbol{A}$ be a $n \times n$–dimensional matrix and let $\boldsymbol{B}$ be a $m \times m$–dimensional matrix. The **Kronecker sum** of $\boldsymbol{A}$ and $\boldsymbol{B}$, denoted by $\boldsymbol{A} \oplus \boldsymbol{B}$, is defined by

$$\boldsymbol{A} \oplus \boldsymbol{B} = \boldsymbol{A} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{B}.$$

For the Kronecker sum, the following results hold.

Proposition A.3 *Let* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *be square matrices. Suppose that* $\mathrm{sp}(A) = \{\lambda_i\}_i$ *and* $\mathrm{sp}(A) = \{\mu_j\}_j$*. Then,*

$$\mathrm{sp}(\boldsymbol{A} \oplus \boldsymbol{B}) = \{\lambda_i + \mu_j\}_{ij}.$$

Proof. See Section 2.4.XIV in Graham (2018).

Proposition A.4 *For any square matrices* $\boldsymbol{A}$ *and* $\boldsymbol{B}$*,*

$$e^{\boldsymbol{A}\oplus\boldsymbol{B}} = e^{\boldsymbol{A}} \otimes e^{\boldsymbol{B}}$$

.

Proof. See Example 2.6 in Graham (2018).

Appendix B

Holomorphic functional calculus

In this section, our aim is to make sense to expresions of the kind $f(\boldsymbol{A})$ where $f : U \subset \mathbb{C} \to \mathbb{C}$ is an holomorphic function and $\boldsymbol{A}$ is a square matrix with complex entries.

A first attempt would be to consider the Laurent series of f about a point $a \in U$, given by

$$f(z) = \sum_{n=1}^{\infty} a_n (z-a)^n$$

with radius of convergence R, and defining

$$f(\boldsymbol{A}) = \sum_{n=1}^{\infty} a_n (\boldsymbol{A} - a\boldsymbol{I})^n, \tag{B.0.1}$$

whenever $\|\boldsymbol{A}\| < R$ for some matrix–norm $\|\cdot\|$. This approach would trivially make matrix–polinomials and the matrix–exponential function consistent with their usual definition. However, other analytic functions would need to be studied via infinite series of power matrices, which is, in many instances, inconvenient. Thus, we seek to work with an equivalent definition of $f(\boldsymbol{A})$, one which is based on the Cauchy integral formula below.

Theorem B.1 (Cauchy integral formula) *Let U be an open subset of $\mathbb{C}$ and let $f : U \to \mathbb{C}$ be a holomorphic function. Let $\gamma \subset U$ be a clock–wise oriented closed chain. Then, for every $a \in \mathbb{C}$ which is winded once by a closed curve of γ we have that*

$$f(a) = \frac{1}{2\pi \mathrm{i}} \oint_{\gamma} \frac{f(z)}{z-a} \mathrm{d}z. \tag{B.0.2}$$

Inspired by this, for matrices $\boldsymbol{A}$ with $\mathrm{sp}(A) \subset U$, we define

$$f(\boldsymbol{A}) = \frac{1}{2\pi \mathrm{i}} \oint_{\gamma} f(z)(z\boldsymbol{I} - \boldsymbol{A})^{-1} \mathrm{d}z,$$

where γ is a closed chain that encloses $\mathrm{sp}(\boldsymbol{A})$, in the sense that each $\lambda \in \mathrm{sp}(\boldsymbol{A})$ is winded once by exactly one closed curve of γ. The previous definition of $f(\boldsymbol{A})$ is known as the Dunford–Riesz functional calculus. In the following, we will make use of the next result.

Theorem B.2 (Resolvent equation) *For $z, \zeta \in \mathbb{C} \setminus \mathrm{sp}(\boldsymbol{A})$,*

$$(z\boldsymbol{I} - \boldsymbol{A})^{-1} - (\zeta\boldsymbol{I} - \boldsymbol{A})^{-1} = (\zeta - z)(z\boldsymbol{I} - \boldsymbol{A})^{-1}(\zeta\boldsymbol{I} - \boldsymbol{A})^{-1} \tag{B.0.3}$$

Proof. Since $(z\boldsymbol{I} - \boldsymbol{A})$ and $(\zeta\boldsymbol{I} - \boldsymbol{A})$ clearly commute, so do $(z\boldsymbol{I} - \boldsymbol{A})^{-1}$ and $(\zeta\boldsymbol{I} - \boldsymbol{A})^{-1}$. Then,

$$\begin{aligned}
&(z\boldsymbol{I} - \boldsymbol{A})^{-1} - (\zeta\boldsymbol{I} - \boldsymbol{A})^{-1} \\
&\quad = (\zeta\boldsymbol{I} - \boldsymbol{A})(\zeta\boldsymbol{I} - \boldsymbol{A})^{-1}(z\boldsymbol{I} - \boldsymbol{A})^{-1} - (z\boldsymbol{I} - \boldsymbol{A})(z\boldsymbol{I} - \boldsymbol{A})^{-1}(\zeta\boldsymbol{I} - \boldsymbol{A})^{-1} \\
&\quad = ((\zeta\boldsymbol{I} - \boldsymbol{A}) - (z\boldsymbol{I} - \boldsymbol{A}))(z\boldsymbol{I} - \boldsymbol{A})^{-1}(\zeta\boldsymbol{I} - \boldsymbol{A})^{-1} \\
&\quad = (\zeta\boldsymbol{I} - z\boldsymbol{I})(z\boldsymbol{I} - \boldsymbol{A})^{-1}(\zeta\boldsymbol{I} - \boldsymbol{A})^{-1},
\end{aligned}$$

and the proof is finished. □

Theorem B.3 *Let f and g be holomorphic functions with domain U, and let $\boldsymbol{A}$ be a square matrix such that $\mathrm{sp}(\boldsymbol{A}) \subset U$. Also, let $a, b \in \mathbb{C}$. Then*

$$a(f(\boldsymbol{A})) + b(g(\boldsymbol{A})) = (af + bg)(\boldsymbol{A}), \quad \textit{and}$$

$$f(\boldsymbol{A})g(\boldsymbol{A}) = (fg)(\boldsymbol{A}).$$

PROOF. The first assertion follows directly by the linearity of integrals. For the second assertion, let $\gamma_f \subset U$ be a closed chain which encloses $\mathrm{sp}(\boldsymbol{A})$ and let $\gamma_g \subset U$ be a closed chain which encloses $\mathrm{sp}(\boldsymbol{A}) \cup \gamma_f$. Then,

$$\begin{aligned}
f(\boldsymbol{A})g(\boldsymbol{A}) &= \left(\frac{1}{2\pi\mathrm{i}}\oint_{\gamma_f} f(z)(z\boldsymbol{I}-\boldsymbol{A})^{-1}\mathrm{d}z\right)\left(\frac{1}{2\pi\mathrm{i}}\oint_{\gamma_g} g(\zeta)(\zeta\boldsymbol{I}-\boldsymbol{A})^{-1}\mathrm{d}\zeta\right)\\
&= \frac{1}{(2\pi\mathrm{i})^2}\oint_{\gamma_f}\oint_{\gamma_g} f(z)g(\zeta)(z\boldsymbol{I}-\boldsymbol{A})^{-1}(\zeta\boldsymbol{I}-\boldsymbol{A})^{-1}\mathrm{d}\zeta\mathrm{d}z\\
&= \frac{1}{(2\pi\mathrm{i})^2}\oint_{\gamma_f}\oint_{\gamma_g} f(z)g(\zeta)\frac{(z\boldsymbol{I}-\boldsymbol{A})^{-1}-(\zeta\boldsymbol{I}-\boldsymbol{A})^{-1}}{\zeta-z}\mathrm{d}\zeta\mathrm{d}z\\
&= \frac{1}{(2\pi\mathrm{i})^2}\oint_{\gamma_f}\left(\oint_{\gamma_g}\frac{g(\zeta)}{\zeta-z}\mathrm{d}\zeta\right) f(z)(z\boldsymbol{I}-\boldsymbol{A})^{-1}\mathrm{d}z\\
&\quad - \frac{1}{(2\pi\mathrm{i})^2}\oint_{\gamma_g}\left(\oint_{\gamma_f}\frac{f(z)}{\zeta-z}\mathrm{d}z\right) g(\zeta)(\zeta\boldsymbol{I}-\boldsymbol{A})^{-1}\mathrm{d}\zeta.
\end{aligned}$$

Since γ_g encloses γ_f, then for every $z \in \gamma_f$ the Cauchy integral formula implies that $\oint_{\gamma_g}\frac{g(\zeta)}{\zeta-z}\mathrm{d}\zeta = (2\pi i)g(z)$. Since γ_f **does not** enclose γ_g, then for every $\zeta \in \gamma_g$, $z \to \frac{f(\zeta)}{\zeta-z}$ is an holomorphic function on the interior of the chain γ_f, implying that $\oint_{\gamma_f}\frac{f(\zeta)}{\zeta-z}\mathrm{d}z = 0$. Thus, the result follows. □

Theorem B.4 *For any square matrix $\boldsymbol{A}$ with $\mathrm{sp}(\boldsymbol{A}) \subset U$, $f(\mathrm{sp}(\boldsymbol{A})) = \mathrm{sp}(f(\boldsymbol{A}))$.*

PROOF. First, let us show that $f(\mathrm{sp}(\boldsymbol{A})) \subseteq \mathrm{sp}(f(\boldsymbol{A}))$. Fix $\lambda \in \mathrm{sp}(\boldsymbol{A})$. Let q be the holomorphic function on U defined by

$$q(z) = \begin{cases} \frac{f(z)-f(\lambda)}{z-\lambda}, & z \neq \lambda \\ f'(z), & z = \lambda. \end{cases}$$

Let γ be a closed chain which encloses $\mathrm{sp}(\boldsymbol{A})$ (and thus encloses λ). Then,

$$\begin{aligned}
f(\boldsymbol{A}) - f(\lambda)\boldsymbol{I} &= \frac{1}{2\pi\mathrm{i}}\oint_{\gamma} g(z) - g(\lambda)(z\boldsymbol{I}-\boldsymbol{A})^{-1}\mathrm{d}z\\
&= \frac{1}{2\pi\mathrm{i}}\oint_{\gamma}(z-\lambda)q(z)(z\boldsymbol{I}-\boldsymbol{A})^{-1}\mathrm{d}z\\
&= (\boldsymbol{A}-\lambda\boldsymbol{I})q(\boldsymbol{A}),
\end{aligned}$$

which implies that,

$$\det(f(\boldsymbol{A}) - f(\lambda)\boldsymbol{I}) = \det(\boldsymbol{A}-\lambda\boldsymbol{I})\cdot\det(q(\boldsymbol{A})) = 0\cdot\det(q(\boldsymbol{A})) = 0.$$

This means that $f(\lambda) \in \mathrm{sp}(f(\boldsymbol{A}))$.

Now, to prove that $\mathrm{sp}(f(\boldsymbol{A})) \subseteq f(\mathrm{sp}(\boldsymbol{A}))$, fix $\mu \in U \setminus f(\mathrm{sp}(\boldsymbol{A}))$. The function $h(z) = (f(z) - \mu)^{-1}$ is holomorphic in a neighbourhood of $\mathrm{sp}(\boldsymbol{A})$, say $U' \subseteq U$. Then, restricted to U', $\boldsymbol{I} = h(\boldsymbol{A})(f(\boldsymbol{A}) - \mu\boldsymbol{I})$, so that $\det(f(\boldsymbol{A}) - \mu\boldsymbol{I}) \neq 0$. Thus, $\mu \notin \mathrm{sp}(f(\boldsymbol{A}))$ and the result is proved. □

Theorem B.5 *Let f be an analytic function defined on $U \subset \mathbb{C}$ and let g be an analytic function defined on $U' \subset \mathbb{C}$. Suppose that* $\mathrm{sp}(A) \subset U'$ *and* $\mathrm{sp}(g(A)) \subset U$. *Then*

$$f(g(\boldsymbol{A})) = (f \circ g)(\boldsymbol{A}).$$

Proof. Let $\gamma_f \subset U$ be a closed chain that encloses $\mathrm{sp}(g(A))$ and let $\gamma_g \subset U'$ be a chain that encloses $\mathrm{sp}(A)$ and is such that $g(\gamma_g) \subset U$: the existance of a chain γ_g that attains the last condition follows by Theorem B.4 and the continuity of g. Then,

$$\begin{aligned}
f(g(\boldsymbol{A})) &= \frac{1}{2\pi\mathrm{i}} \oint_{\gamma_f} f(z)(z\boldsymbol{I} - g(\boldsymbol{A}))^{-1}\mathrm{d}z \\
&= \frac{1}{2\pi\mathrm{i}} \oint_{\gamma_f} f(z) \left(\frac{1}{2\pi i} \oint_{\gamma_g} (z - g(\zeta))^{-1}(\zeta - \boldsymbol{A})^{-1}\mathrm{d}\zeta \right) \mathrm{d}z \\
&= \frac{1}{2\pi\mathrm{i}} \oint_{\gamma_g} \left(\frac{1}{2\pi\mathrm{i}} \oint_{\gamma_f} f(z)(z - g(\zeta))^{-1}\mathrm{d}z \right) (\zeta - \boldsymbol{A})^{-1}\mathrm{d}\zeta \\
&= \frac{1}{2\pi\mathrm{i}} \oint_{\gamma_g} f(g(\zeta))(\zeta - \boldsymbol{A})^{-1}\mathrm{d}\zeta \\
&= (f \circ g)(\boldsymbol{A}),
\end{aligned}$$

and the proof is finished. □

Appendix C

Richardson extrapolation formula

Let $\{f_n\} \subset \mathbb{R}$ be a sequence such that $f_n \to f$ as $n \to \infty$ for some $f \in \mathbb{R}$. Furthermore, suppose that

$$f_n = f + K_1 n^{-1} + K_2 n^{-2} + \cdots + K_m n^{-m} + O(n^{-(m+1)}),$$

for some $m \geq 1$ and (possibly unknown) constants $K_1, K_2, \ldots, K_m$. Define

$$\widehat{f}_n = (n+1) f_{n+1} - n f_n, \quad n \geq 1.$$

Then, it is easily seen that

$$\widehat{f}_n = f + O(n^{-2}).$$

This means that $\widehat{f}_n$ converges to f and does so with an error rate of the order n^{-2}. The sequence $\{\widehat{f}_n\}_{n \geq 1}$ is known as the **Richardson extrapolation of** f. If $m \geq 2$, one can construct approximations that converge to f at a faster rate. For instance, let

$$\mathring{f}_n = \frac{n^2}{2} f_n - (n+1)^2 f_{n+1} + \frac{(n+2)^2}{2} f_{n+2}.$$

Then, it is readily verified that

$$\mathring{f}_n = f + O(n^{-3}),$$

so that $\mathring{f}_n$ converges to f and does so with an error rate of the order n^{-3}. This is somewhat similar to an **iterated** Richardson extrapolation, that is, to apply a Richardson extrapolation to the Richardson extrapolation sequence $\{\widehat{f}_n\}_{n\geq 1}$.

Bibliography

Albrecher, H. and O. J. Boxma (2004). A ruin model with dependence between claim sizes and claim intervals. *Insurance: Mathematics and Economics 35*(2), 245–254.

Albrecher, H., O. J. Boxma, and J. Ivanovs (2014). On Simple Ruin Expressions in Dependent Sparre Andersen Risk Models. *Journal of Applied Probability 51*(01), 293–296.

Albrecher, H. and J. Ivanovs (2017). Strikingly simple identities relating exit problems for Lévy processes under continuous and Poisson observations. *Stochastic Processes and their Applications 127*(2), 643–656.

Aldous, D. and L. Shepp (1987). The least variable phase type distribution is erlang. *Communications in Statistics. Stochastic Models 3*(3), 467–473.

Andersen, E. S. (1957). On the collective theory of risk in case of contagion between claims. *Bulletin of the Institute of Mathematics and its Applications 12*(2), 275–279.

Anick, D., D. Mitra, and M. M. Sondhi (1983). Stochastic Theory of a Data-Handling System with Multiple Sources. *Bell System Technical Journal 61*(8), 1871–1894.

Asmussen, S. (1989). Risk theory in a Markovian environment. *Scandinavian Actuarial Journal 1989*(2), 69–100.

Asmussen, S. (1995a). Stationary distributions for fluid flow models with or without brownian noise. *Communications in Statistics. Stochastic Models 11*(1), 21–49.

Asmussen, S. (1995b). Stationary distributions via first passage times. In *Advances in Queueing Theory, Methods, and Open Problems*, pp. 79–102. CRC Press LLC.

Asmussen, S. (2003). *Applied Probability and Queues*, Volume 51 of *Stochastic Modelling and Applied Probability*. New York, NY: Springer New York.

Asmussen, S. and H. Albrecher (2010). *Ruin Probabilities*, Volume 14 of *Second Edition*. World Scientific Publishing Co. Pte. Ltd.

Asmussen, S., F. Avram, and M. R. Pistorius (2004). Russian and American put options under exponential phase-type Lévy models. *Stochastic Processes and their Applications 109*(1), 79–111.

Asmussen, S., F. Avram, and M. Usabel (2002). Erlangian Approximations for Finite-Horizon Ruin Probabilities. *ASTIN Bulletin 32*(02), 267–281.

Asmussen, S. and M. Bladt (1999). Point processes with finite-dimensional conditional probabilities. *Stochastic Processes and their Applications 82*(1), 127–142.

Asmussen, S. and J. Ivanovs (2018). A factorization of a Lévy process over a phase-type horizon. *arXiv preprint arXiv:1803.00273*.

Asmussen, S. and G. Koole (1993). Marked point processes as limits of Markovian arrival streams. *Journal of Applied Probability 30*(02), 365–372.

Asmussen, S., O. Nerman, and M. Olsson (1996). Fitting phase-type distributions via the EM algorithm. *Scandinavian Journal of Statistics*, 419–441.

Asmussen, S. and C. A. O'Cinneide (2006). *Matrix-Exponential Distributions*. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Asmussen, S. and D. Perry (1998). An Operational Calculus for Matrix-Exponential Distributions, with Applications to a Brownian (q, Q) Inventory Model. *Mathematics of Operations Research 23*(1), 166–176.

Asmussen, S. and T. Rolski (1992). Computational methods in risk theory: A matrix-algorithmic approach. *Insurance: Mathematics and Economics 10*(4), 259–274.

Assaf, D., N. A. Langberg, T. H. Savits, and M. Shaked (1984). Multivariate Phase-Type Distributions. *Operations Research 32*(3), 688–702.

Avram, F., A. L. Badescu, M. R. Pistorius, and L. Rabehasaina (2016). On a class of dependent Sparre Andersen risk models and a bailout application. *Insurance: Mathematics and Economics 71*, 27–39.

Badescu, A., L. Breuer, A. Da Silva Soares, G. Latouche, M.-A. Remiche, and D. Stanford (2005). Risk processes analyzed as fluid queues. *Scandinavian Actuarial Journal 2005*(2), 127–141.

Badescu, A. L., E. C. K. Cheung, and D. Landriault (2009). Dependent Risk Models with Bivariate Phase-Type Distributions. *Journal of Applied Probability 46*(01), 113–131.

Badila, E. S., O. J. Boxma, and J. A. C. Resing (2014). Queues and Risk Processes with Dependencies. *Stochastic Models 30*(3), 390–419.

Badila, E. S., O. J. Boxma, and J. A. C. Resing (2015). Two parallel insurance lines with simultaneous arrivals and risks correlated with inter-arrival times. *Insurance: Mathematics and Economics 61*, 48–61.

Baker, R. (2008). An order-statistics-based method for constructing multivariate distributions with fixed marginals. *Journal of Multivariate Analysis 99*(10), 2312–2327.

Baurdoux, E. J., J. C. Pardo, J. L. Pérez, and J.-F. Renaud (2016). Gerber–Shiu distribution at Parisian ruin for Lévy insurance risk processes. *Journal of Applied Probability 53*(02), 572–584.

Bean, N., G. T. Nguyen, and F. Poloni (2018). Doubling algorithms for stationary distributions of fluid queues: A probabilistic interpretation. *Performance Evaluation 125*, 1–20.

Bean, N. G. and B. F. Nielsen (2010). Quasi-Birth-and-Death Processes with Rational Arrival Process Components. *Stochastic Models 26*(3), 309–334.

Bean, N. G., M. M. O'Reilly, and P. G. Taylor (2005). Algorithms for return probabilities for stochastic fluid flows . *Stochastic Models 21*(1), 149–184.

Bertoin, J. (1998). *Lévy processes*, Volume 121. Cambridge University Press.

Bladt, M. et al. (1996). Renewal theory and queueing algorithms for matrix-exponential distributions. In *Matrix-analytic Methods in Stochastic Models*, pp. 313–341. Marcel Dekker Incorporated.

Bladt, M., A. C. Navarro, and B. F. Nielsen (2015). On the use of functional calculus for phase-type and related distributions. *Stochastic Models 32*(1), 1–19.

Bladt, M. and M. F. Neuts (2003). Matrix-Exponential Distributions: Calculus and Interpretations via Flows. *Stochastic Models 19*(1), 113–124.

Bladt, M. and B. F. Nielsen (2010a). Multivariate Matrix-Exponential Distributions. *Stochastic Models 26*(1), 1–26.

Bladt, M. and B. F. Nielsen (2010b). On the Construction of Bivariate Exponential Distributions with an Arbitrary Correlation Coefficient. *Stochastic Models 26*(2), 295–308.

Bladt, M. and B. F. Nielsen (2017). *Matrix-Exponential Distributions in Applied Probability*, Volume 81 of *Probability Theory and Stochastic Modelling*. Boston, MA: Springer US.

Bladt, M., B. F. Nielsen, and O. Peralta (2018). Parisian types of ruin probabilities for a class of dependent risk-reserve processes. *Scandinavian Actuarial Journal*, 1–30.

Bladt, M., B. F. Nielsen, and G. Samorodnitsky (2014). Calculation of ruin probabilities for a dense class of heavy tailed distributions. *Scandinavian Actuarial Journal 2015*(7), 573–591.

Borst, S. C., O. J. Boxma, and M. B. Combé (1992). Collection of customers. *ACM SIGMETRICS Performance Evaluation Review 20*(1), 47–59.

Boudreault, M., H. Cossette, D. Landriault, and E. Marceau (2006). On a risk model with dependence between interclaim arrivals and claim sizes. *Scandinavian Actuarial Journal 2006*(5), 265–285.

Breuer, L. (2008). First Passage Times for Markov Additive Processes with Positive Jumps of Phase Type. *Journal of Applied Probability 45*(03), 779–799.

Breuer, L. (2010). A quintuple law for Markov additive processes with phase-type jumps. *Journal of Applied Probability 47*(02), 441–458.

Breuer, L. (2016). A semi-explicit density function for Kulkarni's bivariate phase-type distribution. *Stochastic Models 32*(4), 632–642.

Buchholz, P. and M. Telek (2010). Stochastic Petri nets with matrix exponentially distributed firing times. *Performance Evaluation 67*(12), 1373–1385.

Buchholz, P. and M. Telek (2011). On minimal representations of Rational Arrival Processes. *Annals of Operations Research 202*(1), 35–58.

Buchholz, P. and M. Telek (2012). Rational Processes Related to Communicating Markov Processes. *Journal of Applied Probability 49*(01), 40–59.

Buchholz, P. and M. Telek (2013). Rational Automata Networks: A Non-Markovian Modeling Approach. *INFORMS Journal on Computing 25*(1), 87–101.

Conolly, B. W. (1968). The Waiting Time Process for a Certain Correlated Queue. *Operations Research 16*(5), 1006–1015.

Cox, D. R. (1955a). A use of complex probabilities in the theory of stochastic processes. *Mathematical Proceedings of the Cambridge Philosophical Society 51*(02), 313–319.

Cox, D. R. (1955b). The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables. *Mathematical Proceedings of the Cambridge Philosophical Society 51*(03), 433.

Cramér, H. (1930). On the Mathematical Theory of Risk. In *Collected Works II*, pp. 601–678. Berlin, Heidelberg: Springer Berlin Heidelberg.

Czarna, I., Y. Li, Z. Palmowski, and C. Zhao (2017). The joint distribution of the Parisian ruin time and the number of claims until Parisian ruin in the classical risk model. *Journal of Computational and Applied Mathematics 313*, 499–514.

Czarna, I. and Z. Palmowski (2011). Ruin Probability with Parisian Delay for a Spectrally Negative Lévy Risk Process. *Journal of Applied Probability 48*(04), 984–1002.

Da Silva Soares, A. and G. Latouche (2002). Further results on the similarity between fluid queues and QBDs. In *Proceedings of the Fourth International Conference*, pp. 89–106. WORLD SCIENTIFIC.

Dassios, A. and S. Wu (2008). Parisian ruin with exponential claims. Unpublished manuscript. Available at http://stats.lse.ac.uk/angelos/docs/exponentialjump.pdf.

Davis, M. H. (1984). Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 353–388.

Davis, M. H. A. (1976). The Representation of Martingales of Jump Processes. *SIAM Journal on Control and Optimization 14*(4), 623–638.

Davis, M. H. A. (1993). *Markov Models and Optimization*. Boston, MA: Springer US.

Doolittle, E. (1998). Analytic Functions of Matrices. Technical report. Available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.51.2968&rep=rep1&type=pdf.

Dou, X., S. Kuriki, and G. D. Lin (2013). Dependence structures and asymptotic properties of Baker's distributions with fixed marginals. *Journal of Statistical Planning and Inference 143*(8), 1343–1354.

Dou, X., S. Kuriki, G. D. Lin, and D. Richards (2016). EM algorithms for estimating the Bernstein copula. *Computational Statistics & Data Analysis 93*, 228–245.

Dunford, N. and J. T. Schwartz (1958). *Linear Operators: General theory*. Pure and applied mathematics. Interscience Publishers.

Dynkin, E. B. (1965). *Markov Processes*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Erlang, A. K. (1909). The theory of probabilities and telephone conversations. *Nyt. Tidsskr. Mat. Ser. B 20*, 33–39.

Fackrell, M. W. (2003). *Characterization of matrix-exponential distributions*. Ph. D. thesis.

Graham, A. (2018). *Kronecker products and matrix calculus with applications*. Courier Dover Publications.

Grandell, J. (1977). Point processes and random measures. *Advances in Applied Probability 9*(03), 502–526.

Guérin, H. and J.-F. Renaud (2017). On the distribution of cumulative Parisian ruin. *Insurance: Mathematics and Economics 73*, 116–123.

He, Q.-M., H. Zhang, and J. C. Vera (2012). On Some Properties of Bivariate Exponential Distributions. *Stochastic Models 28*(2), 187–206.

Hirsch, M., S. Smale, and R. Devaney (2013). *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Elsevier.

Horváth, I., O. Sáfár, M. Telek, and B. Zámbó (2016). Concentrated Matrix Exponential Distributions. In *Computer Performance Engineering*, pp. 18–31. Cham: Springer International Publishing.

Horváth, I. and M. Telek (2015). A Constructive Proof of the Phase-Type Characterization Theorem. *Stochastic Models 31*(2), 316–350.

Itô, K. (1941). On stochastic processes (I). In *Japanese journal of mathematics: transactions and abstracts*, pp. 261–301. The Mathematical Society of Japan.

Ivanovs, J. (2011). *One-sided Markov additive processes and related exit problems*. Ph. D. thesis, Universiteit van Amsterdam.

Jacobsen, M. (2006). *Point Process Theory and Applications: Marked Point and Piecewise Deterministic Processes*. Probability and Its Applications. Birkhäuser Boston.

Jagerman, D. L. (1982). An Inversion Technique for the Laplace Transform. *Bell System Technical Journal 61*(8), 1995–2002.

Jensen, A. (1954). *A Distribution Model, Applicable to Economics*. Copenhagen.

Karandikar, R. L. and V. G. Kulkarni (1995). Second-Order Fluid Flow Models: Reflected Brownian Motion in a Random Environment. *Operations Research 43*(1), 77–88.

Khinchin, A. Y. (1967). The mathematical theory of a stationary queue. Technical report.

Khintchine, A. (1937). A new derivation of one formula by Lévy P. *Bulletin of Moscow State University 1*(1), 1.

Kulkarni, V. G. (1989). A New Class of Multivariate Phase Type Distributions. *Operations Research 37*(1), 151–158.

Kuznetsov, A., A. E. Kyprianou, and V. Rivero (2012). The Theory of Scale Functions for Spectrally Negative Lévy Processes. In *Lévy Matters II*, pp. 97–186. Berlin, Heidelberg: Springer Berlin Heidelberg.

Kyprianou, A. E. (2014). *Fluctuations of Lévy Processes with Applications*. Universitext. Berlin, Heidelberg: Springer Berlin Heidelberg.

Landriault, D., J.-F. Renaud, and X. Zhou (2011). Occupation times of spectrally negative Lévy processes with applications. *Stochastic Processes and their Applications 121*(11), 2629–2641.

Landriault, D., J.-F. Renaud, and X. Zhou (2013). An Insurance Risk Model with Parisian Implementation Delays. *Methodology and Computing in Applied Probability 16*(3), 583–607.

Latouche, G. and G. Nguyen (2018). Analysis of fluid flow models. *arXiv preprint arXiv:1802.04355*.

Lévy, P. (1934). Sur les intégrales dont les éléments sont des variables aléatoires indépendantes. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze 3*(3-4), 337–366.

Loeffen, R., I. Czarna, and Z. Palmowski (2013). Parisian ruin probability for spectrally negative Lévy processes. *Bernoulli 19*(2), 599–609.

Loynes, R. M. (1961). A continuous-time treatment of certain queues and infinite dams. *Journal of the Australian Mathematical Society 2*(04), 484–498.

Maier, R. S. (1991). The algebraic construction of phase-type distributions. *Communications in Statistics. Stochastic Models 7*(4), 573–602.

Marshall, A. W., I. Olkin, and B. C. Arnold (2011). *Inequalities: Theory of Majorization and Its Applications*. Springer Series in Statistics. New York, NY: Springer New York.

Mitchell, K. (2001). Constructing a correlated sequence of matrix exponentials with invariant first-order properties. *Operations Research Letters 28*(1), 27–34.

Mocanu, Ş. and C. Commault (1999). Sparse representations of phase-type distributions. *Communications in Statistics. Stochastic Models 15*(4), 759–778.

Müller, A. (2000). On the waiting times in queues with dependency between interarrival and service times. *Operations Research Letters 26*(1), 43–47.

Neuts, M. F. (1979). A versatile Markovian point process. *Journal of Applied Probability 16*(04), 764–779.

Neuts, M. F. (1994). *Matrix-geometric solutions in stochastic models: an algorithmic approach*. Courier Corporation.

Nguyen, G. T. and F. Poloni (2014). Componentwise accurate fluid queue computations using doubling algorithms. *Numerische Mathematik 130*(4), 763–792.

Nguyen, G. T. and F. Poloni (2016). Componentwise accurate brownian motion computations using cyclic reduction. *arXiv preprint arXiv:1605.01482*.

O'Cinneide, C. A. (1990). Characterization of phase-type distributions. *Communications in Statistics. Stochastic Models 6*(1), 1–57.

O'Cinneide, C. A. (1991a). Phase-type distributions and invariant polytopes. *Advances in Applied Probability 23*(03), 515–535.

O'Cinneide, C. A. (1991b). Phase-Type Distributions and Majorization. *The Annals of Applied Probability 1*(2), 219–227.

Panjer, H. H. (1986). Direct Calculation of Ruin Probabilities. *The Journal of Risk and Insurance 53*(3), 521.

Peralta, O., L. Rojas-Nandayapa, W. Xie, and H. Yao (2018). Approximation of ruin probabilities via Erlangized scale mixtures. *Insurance: Mathematics and Economics 78*, 136–156.

Pollaczek, F. (1930). Über eine Aufgabe der Wahrscheinlichkeitstheorie. I. *Mathematische Zeitschrift 32*(1), 64–100.

Ramaswami, V. (1999). Matrix analytic methods for stochastic fluid flows. In *Proc. of ITC 16*. Elsevier.

Ramaswami, V. (2006). Passage Times in Fluid Models with Application to Risk Processes. *Methodology and Computing in Applied Probability 8*(4), 497–515.

Ramaswami, V., D. G. Woolford, and D. A. Stanford (2008). The erlangization method for Markovian fluid flows. *Annals of Operations Research 160*(1), 215–225.

Ren, J., L. Breuer, D. A. Stanford, and K. Yu (2009). Perturbed Risk Processes Analyzed as Fluid Flows. *Stochastic Models 25*(3), 522–544.

Rogers, L. C. G. (1994). Fluid Models in Queueing Theory and Wiener-Hopf Factorization of Markov Chains. *The Annals of Applied Probability 4*(2), 390–413.

Rogers, L. C. G. and D. Williams (1993). *Diffusions, Markov Processes and Martingales* (2 ed.). Cambridge: Cambridge University Press.

Rogozin, B. A. (1966). On Distributions of Functionals Related to Boundary Problems for Processes with Independent Increments. *Theory of Probability & Its Applications 11*(4), 580–591.

Rolski, T., H. Schmidli, V. Schmidt, and J. L. Teugels (2009). *Stochastic Processes for Insurance and Finance*. Wiley Series in Probability and Statistics. Wiley.

Rudin, W. (2006). *Real and complex analysis*. Tata McGraw-Hill Education.

Rudnicki, R. and M. Tyran-Kamińska (2017). *Piecewise Deterministic Processes in Biological Models*. SpringerBriefs in Applied Sciences and Technology. Cham: Springer International Publishing.

Sancetta, A. and S. Satchell (2004). The Bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric Theory 20*(03), 600.

Sato, K.-i. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge University Press.

Simon, M. (2017). *Markov-modulated processes: Brownian motions, option pricing and epidemics*. Ph. D. thesis, Université Libre de Bruxelles.

Spitzer, F. (1957). The Wiener-Hopf equation whose kernel is a probability density. *Duke Mathematical Journal 24*(3), 327–343.

Stanford, D. A., F. Avram, A. L. Badescu, L. Breuer, A. D. Silva Soares, and G. Latouche (2005). Phase-type Approximations to Finite-time Ruin Probabilities in the Sparre-Andersen and Stationary Renewal Risk Models. *ASTIN Bulletin 35*(01), 131–144.

Telek, M. (2011). Some Structural Properties of Markov and Rational Arrival Processes. *Stochastic Models 27*(4), 723–746.

Vatamidou, E., I. J. B. F. Adan, M. Vlasiou, and B. Zwart (2012). On the accuracy of phase-type approximations of heavy-tailed risk models. *Scandinavian Actuarial Journal 2014*(6), 510–534.

Wiener, N. and R. Paley (1934). *Fourier Transforms in the Complex Domain*, Volume 19 of *Colloquium Publications*. Providence, Rhode Island: American Mathematical Society.